Bildung zwischen Unsicherheit und Evidenz

10 GEBF

Zehnte Tagung der Gesellschaft für Empirische Bildungsforschung

Universität Duisburg-Essen, Essen

Nachwuchstagung
27.02.2023

Haupttagung
28.02. – 02.03.2023

ABSTRACTBAND

GEBF Gesellschaft für Empirische Bildungsforschung

izfb Interdisziplinäres Zentrum für Bildungsforschung

Zentrum für Lehrkräftebildung ZfB

dokFORUM

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Inhalt

Inhalt ... 1
Keynotes ... 1
Keynote 1: Prof. Dr. Gerd Gigerenzer ... 1
Keynote 2: Prof. Dr. Eckhard Klieme, Preisträger des GEBF-Preises für das Lebenswerk ... 1
Keynote 3: Prof. Dr. Monika Waldis Weber ... 2
Keynote 4: Prof. Dr. Julika Griem ... 3
Keynote Nachwuchstagung: Prof. Dr. Matthias Brand ... 3
Session A ... 4
Einzelbeiträge ... 4
A-01: Krisen als Unterrichtsinhalte ... 4
A-02: Feedback und Kollaboration ... 8
A-03: Struktur Bildungssystem ... 15
A-04: Unterrichtsmerkmale ... 22
A-05: Sprache im Unterricht ... 29
A-06: Berufswahl ... 35
A-07: Erwartungs-Wert-Theorien ... 40
A-08: Heterogenität Grundschule ... 47
A-09: Schulleitung und Schulaufsicht ... 52
A-10: Gendergerechtigkeit ... 58
Symposien ... 64
A-11: Die Entwicklung von Schüler:innen im Verlauf der COVID-19-Pandemie – Ein Stand 3 Jahre nach Pandemiebeginn ... 64
A-12: Motivation im Unterrichtskontext ... 70
A-13: Unterricht während der Pandemie: Was wissen wir über die Qualität des Distanzunterrichts? ... 77
A-14: Understanding and decreasing discrimination in school contexts ... 83
A-15: Simulationsbasiertes Lernen und Messen in der Mediziner:innen und Lehrer:innenausbildung ... 88
A-16: Ursachen und Wirkungen irregulärer Bildungsverläufe: Empirische Befunde aus Deutschland und Luxembourg ... 96
A-17: Metakognitive Überwachung beim Lernen in Schule und Hochschule: Maßnahmen zur Förderung der Urteilsgenauigkeit ... 102
A-18: Professionelle Kompetenz von Fachkräften im Elementarbereich – Gemeinsamkeiten und Spezifika in den Bildungsbereichen Mathematik, Naturwissenschaften und Technik ... 109
A-19: Medienbezogene Lehrkompetenz: Messung, Wirkung, Förderung ... 115
A-20: Zusammenhänge zwischen der Kompetenz von Lehrkräften, Selbstregulation zu fördern, und der Selbstregulation ihrer Lernenden ... 123
Session B ... 130
Einzelbeiträge ... 130
B-01: Kooperation mit Eltern ... 130
B-02: KI und online Lernen ... 136
B-03: COVID-19 Pandemie I ... 143

B-04: Diagnostik und Mathematik ... 150
B-05: Grundschule ... 157
B-06: Diagnostik durch Lehrkräfte ... 162
B-07: Sprachsensibilität ... 169
B-08: Vor- und Grundschule ... 176
B-09: Studienwahl und Studienmotivation ... 183
B-10: Akademische Professionalisierung ... 188
Symposien ... 194
B-11: Zur quantitativen Vermessung kultureller Bildung – methodologische Reflexionen zu Datenverfügbarkeit und Operationalisierung am Beispiel aktueller Forschungsansätze ... 194
B-12: Determinanten des subjektivem Wohlbefindens im Kinder- und Jugendalter ... 202
B-13: Herausforderungen und Entwicklungen bei der Messung schriftsprachlicher Fähigkeiten ... 210
B-14: Professionalisierung datengestützter Entscheidungen für Schule und Unterricht ... 217
B-15: Forschungsdaten erheben, sichern, weitergeben und nachnutzen – Praxisbeispiele aus dem Verbund Forschungsdaten Bildung ... 225
B-16: Designing effective digital learning environments ... 229
B-17: Phasen digital-gestützten individualisierten Übens im Fokus: Erkenntnisse aus dem CODIP-Projekt ... 235
B-18: Unterrichtserleben in Virtual Reality als Chance für die Lehrkräftebildung ... 242
B-19: Emotionen und Emotionsregulation von Lehramtsstudierenden und deren Bedeutung für Professionalisierungsprozesse ... 250
B-20: Was tun? Perspektiven für eine Unterrichtsqualitätsforschung der Zukunft ... 258
Session C ... 264
Einzelbeiträge ... 264
C-01: Digitales Lernen I ... 264
C-02: Einfluss der Eltern ... 269
C-03: Kompetenzentwicklung und -messung ... 275
C-04: Außerunterrichtliche Angebote ... 281
C-05: Unterrichtsqualität ... 287
C-06: Bildungshintergrund und Bildungserfolg ... 293
C-07: Hochschulbildung ... 299
C-08: Schulentwicklung ... 306
C-09: Digitales Lernen II ... 312
C-10: Lehrkräfte und e-Learning ... 319
Symposien ... 325
C-11: Individuelle und kontextuelle Einflussfaktoren der sozial-emotionalen Entwicklung in der Kindheit ... 325
C-12: Digitale Transformation von Schulen der Sekundarstufe II in Deutschland und der Schweiz aus Schulentwicklungsperspektive ... 331
C-13: Evidence-based teacher education – Effekte von evidenzbasierten Lehr-Lern-Prinzipien auf die Wirksamkeit videobasierter Lernumgebungen in der Lehrkräftebildung ... 337
C-14: Psychische Gesundheit von Schüler:innen als Dimension inklusiver Schule: Beiträge aus der Lehrkräfteprofessionalisierungsforschung ... 344

C-15: Trainingsprogramme zur sprachlichen Förderung im Unterricht: Möglichkeiten und Grenzen ... 351
C-16: Simulationsbasiertes Lernen an der Hochschule – neue Erkenntnisse aus experimentellen und metaanalytischen Studien ... 357
C-17: Der Umgang von Lehrkräften mit Herausforderungen in der Bildung für nachhaltige Entwicklung (BNE) ... 365
C-18: Stereotype im Lehrer*innenzimmer und ihre Auswirkungen auf das heterogene Klassenzimmer ... 370
C-19: School adjustment and success of recently arrived refugee and immigrant children in Germany ... 377
C-20: Eigenvideografie als Professionalisierungstool für Lehramtsstudierende im Praxissemester ... 382
Sonderveranstaltung ... 389
C-21: Sonderveranstaltung: DFG ... 389
Session D ... 390
Einzelbeiträge ... 390
D-01: COVID-19 Pandemie II ... 390
D-02: Evidenz und Kredibilität ... 396
D-03: Wissenschaftskommunikation ... 402
D-04: Lehrkräftebildung ... 409
D-05: Inklusive Schule: Gemeinsamer Unterricht ... 415
D-06: Diversität: Schüler*innenleistungen und Ambitionen ... 421
D-07: Förderung und Unterstützung ... 427
D-08: Belastung von Lehrkräften ... 433
D-09: Unterrichtsqualität II ... 439
D-10: Motivation und Interesse ... 446
Symposien ... 452
D-11: Effektivität von Rater*innentrainings: Wie gelingt die Beobachtung der Unterrichtsqualität durch externe Beobachtende? ... 452
D-12: Internationale Studierende in Deutschland – Interdisziplinäre Befunde zu Studienerleben, Studienerfolg und Verbleibsintentionen ... 459
D-13: Situationale Emotionen und Motivation von Studierenden ... 469
D-14: Schulischer Wandel in der Migrationsgesellschaft – Schulkultur(en) im Kontext aktueller Fluchtmigration ... 475
D-15: Bedarfsorientierte Ressourcensteuerung im Mehrebenensystem – durch differenzierte Mittelzuweisung Bildungsungleichheiten verringern? ... 481
D-16: Digitale Kompetenzen und Lehrerbildung ... 488
D-17: Fünf Jahre nach der Erfassung der ‚Global Competences' in PISA: Entwicklungsperspektiven für zukünftige Erhebungen ... 494
D-18: Evaluation digitaler Lernumgebungen zur Förderung der Klassenführungskompetenz bei Lehramtsstudierenden ... 500
D-19: Fachlich-interdisziplinäre Perspektiven auf das Textverstehen: Studien zu Determinanten und Verstehensprozessen im Umgang mit Fachtexten und Aufgaben ... 508
D-20: Die Bedeutung von aktiver Schüler*innenbeteiligung im Unterricht: Individuelle Lernvoraussetzungen, kontextuelle Determinanten und der Zusammenhang mit Lernerfolg ... 515
Sondereranstaltung ... 523

D-21: Sonderveranstaltung: RatSWD ... 523
Session E ... 524
Einzelbeiträge ... 524
E-01: Soziologische Fragen ... 524
E-02: Fachdidaktisches Wissen ... 530
E-03: Professionalisierung von Lehrkräften ... 537
E-04: Entwicklung von Testinstrumenten ... 542
E-05: Familiärer Hintergrund ... 548
E-06: Lernen, Kognition und Enthusiasmus ... 554
E-07: Fachliche Lernpfade ... 559
E-08: Berufs- und Karriereentscheidungen ... 566
E-09: Beraten und Beurteilen ... 573
E-10: Überzeugungen von Lehrkräften ... 579
Symposien ... 585
E-11: Lernen auf Abruf? Balancieren zwischen Anstrengung und Abruferfolg ... 585
E-12: Von der Konzeptualisierung zur Messung: Instrumente zur Erfassung politischen Wissens und deren Anwendung in der empirischen Bildungsforschung ... 593
E-13: A Multi-Faceted Perspective on Current Approaches to Reflect on Teacher Reflections 600
E-14: Potentiale von Log-Daten in authentischen computerbasierten Assessments ... 606
E-15: Forschung und Praxis gemeinsam denken: Wie gelingt das Zusammenwirken von Bildungsforschung und Bildungspraxis? ... 614
E-16: (Virtuelles) kooperatives Arbeiten und Lernen ... 622
E-17: Eye Tracking als Methode zur Erfassung professioneller Unterrichtswahrnehmung ... 628
E-18: Lehrer-Schüler-Beziehungen: Von der Generalisierbarkeit positiver Befunde ... 636
E-19: Bedeutung von Online-Self-Assessments für Studienwahl und Studienerfolg ... 644
E-20: Wohlbefinden und Emotionen von Kindern und Jugendlichen in Zeiten der Corona-Pandemie ... 651
Sonderveranstaltung ... 658
E-21: Sonderveranstaltung: Esther Wojcicki ... 658
Session F ... 659
Einzelbeiträge ... 659
F-01: Zuwanderung und kulturelle Diversität ... 659
F-02: Digitales Lernen III ... 665
F-03: Unterrichtsqualität, Instruktionsdesign und epistemische Emotionen ... 670
F-04: Bildungserfolg und -verlauf: Profile und Faktoren ... 676
F-05: Motivation, Selbstwirksamkeit, Selbstkonzept und Engagement ... 682
F-06: Berufsschulunterricht gestalten ... 689
F-07: Teilhabe und deren Ermöglichung ... 696
F-08: Einstellungen und Überzeugungen von Lehrkräften ... 702
F-09: Wohlbefinden von Lehrkräften ... 707
Symposien ... 714
F-10: Einfluss von Lehrkraftvariablen auf Schüler:innen in heterogenen Lerngruppen bzw. in inklusiven Klassen ... 714

F-11: Adaptive Lernunterstützung im mathematischen, sprachlichen und kognitiven Förderbereich in der Kita ... 720
F-12: Trends im nationalen und internationalen Large-Scale-Assessment. Schulische Bildung in bewegten Zeiten ... 727
F-13: Potenziale geskripteter Unterrichtsvideos für die Lehrkräftebildung ... 734
F-14: Teacher Noticing - Konzepte, Messung, Ergebnisse und Perspektiven ... 741
F-15: Kompetenzentwicklung in der beruflichen Ausbildung: Prädiktoren und Outcomes differenzieller Lernverläufe ... 749
F-16: Schulleitung in Zeiten der Krise: Heraus- oder doch Überforderung? ... 757
F-17: Musikalische Entwicklung im Kindes- und Jugendalter im Zusammenspiel mit kognitiven, motivationalen und psychosozialen Faktoren ... 764
F-18: Motivation und Emotion von Studierenden im Hochschulkontext – Entwicklungstendenzen, Auswirkungen und Förderstrategien einer positiven akademischen Entwicklung ... 770
F-19: Survey-Methoden in der Hochschulforschung ... 778
F-20: Evidenz in der empirischen Bildungsforschung: Auch eine Frage der Methode? ... 785
F-21: Qualitätsmerkmale und Wirksamkeit von Erklärvideos ... 793
Session G ... 800
Einzelbeiträge ... 800
G-01: Persönlichkeit und Kompetenz in verschiedenen Domänen ... 800
G-02: Beziehungen und Strukturen ... 806
G-03: Mathematikunterricht ... 812
G-04: Instruktionale Methoden ... 818
G-05: Prokrastination und Resilienz ... 824
G-06: Unterrichtsgestaltung: Authentisch, individualisiert und bedürfnisorientiert ... 831
G-07: Lehrmaterial und Training ... 838
G-08: Feedback in der Hochschule ... 845
G-09: Motivation und Interesse im Studium ... 852
G-10: Expertise und Überzeugungen von Lehrkräften ... 858
Symposien ... 865
G-11: Von Wissen zu Performanz von Lehrkräften: Beiträge zur Erfassung von Kompetenzfacetten im integrativen Kompetenzverständnis ... 865
G-12: Mehr Studienmotivation in die Hochschulen: Wertüberzeugungen von Studierenden und ihre Förderung in spezifischen Studienkontexten ... 872
G-13: Wirksamkeit und Transferpotentiale von Berufsorientierungsmaßnahmen ... 879
G-14: Stereotype von (angehenden) Lehrkräften ... 885
G-15: Monitoring nachgeordneter Bildungsbereiche in Zeiten gesellschaftlicher Unsicherheit. Ansätze und Instrumente zur Erfassung der politischen und kulturellen Bildung sowie zur Erfassung der Bildung für nachhaltige Entwicklung ... 893
G-16: Kognitive Aktivierung ohne Adaptivität? Notwendigkeit, Möglichkeiten und Herausforderungen einer Verknüpfung beider Konzepte ... 901
G-17: Förderung von Lesekompetenz – What works? ... 908
G-18: Die Professionelle Wahrnehmung von Lehrkräften fachspezifisch messen und fördern: Befunde aus vier Fächern ... 916
G-19: „Bedrohung“ der Testwertinterpretation durch Testmotivation? – Fragen der Validität in der Kompetenzforschung ... 924

G-20: Neue Perspektiven auf das Akademische Selbstkonzept ... 931
Session H ... 939
Einzelbeiträge ... 939
H-01: Professionswissen von Lehrkräften ... 939
H-02: Motivation, Selbstkonzept, Selbstwirklichkeit, leistungszielorientiert und als Vergleich ... 946
H-03: Metakognition und Feedback ... 952
H-04: Herausforderungen in der Hochschulbildung ... 958
H-05: Evidenz, Klausurnoten und Ausbildungserfahrung ... 965
H-06: Selbstreguliertes Lernen ... 971
H-07: Abitur und automatisches Scoring ... 977
H-08: Wissen und Stress bei Lehrkräften ... 983
H-09: Cheating und automatisches Scoring ... 990
H-10: Lernen in MINT-Fächern ... 996
Symposien ... 1003
H-11: Qualität und Wirkung von Mathematikunterricht – Innovative Forschungsansätze auf der Basis der TALIS-Videostudie ... 1003
H-13: Fachspezifische Lehrkräftekompetenzen im Erklären: Überblick und Ergebnisse der drei Projekte FALKE-q, -e, -d ... 1011
H-15: Unterrichtsplanungskompetenz angehender Lehrkräfte: Planungsentscheidungen erfassen, analysieren und reflektieren ... 1018
H-16: Bewertung von Schüler*innentexten - von Menschen und Maschinen ... 1026
H-17: Die Bedeutung der beruflichen Orientierung für einen erfolgreichen Übergang von der Schule in die Ausbildung ... 1034
H-18: Förderbezogene Diagnostik im inklusiven Mathematikunterricht ... 1041
H-19: Informelles Lernen von Lehrpersonen in Deutschland – Analoge und digitale Perspektiven der Professionalisierung ... 1047
H-20: Feedback als Faktor für Lernerfolg und Unterrichtsentwicklung – welche Befunde liefern Interventionsstudien? ... 1053
Postersession ... 1061

Keynotes

Keynote 1: Prof. Dr. Gerd Gigerenzer

Zeit: Dienstag, 28.02.2023: 10:00 - 12:00 · *Ort:* R14 R00 A04

Risikokompetenz: Bildung im digitalen Zeitalter

Gerd Gigerenzer
Max-Planck-Institut für Bildungsforschung

Digitale Risikokompetenz ist die Fähigkeit, die Chancen und Risiken von digitalen Technologien zu verstehen und auch entschlossen zu sein, in einer von Algorithmen bevölkerten Welt die Kontrolle zu behalten. Beispielsweise werden weltweit Milliarden investiert, um Tablets und Whiteboards in Klassenzimmer zu bringen, aber kaum etwas in die Risikokompetenz der Jugendlichen. Folglich zeigen Studien in den USA und Deutschland, dass die Mehrzahl der Digital Natives nicht weiß, wie man die Vertrauenswürdigkeit einer Webseite beurteilt, Fakten von Meinungen unterscheidet, oder auch nur Nachrichten von Werbung. Ob Tablets die Leistung der Schüler verbessern, ist dagegen empirisch umstritten. Auf der gesellschaftlichen Ebene versuchen Tech-Unternehmen und Staaten zunehmend das Verhalten der Bürger vorherzusagen und zu steuern, etwa mittels Big-Nudging Wahlen zu beeinflussen oder mittels sozialer Kreditsysteme autokratische Regierungen zu festigen. In diesem Vortrag werde ich argumentieren, dass der derzeitige Fokus der Bildungspolitik auf die Bereitstellung von Geräten und schnellem Internet sich verändern muss, und zwar hin auf die Ausbildung der Risikokompetenz von Lehrern und Schülern. Wir brauchen mehr Programme welche digitale Risikokompetenz vermitteln, einschließlich (i) dem Erlernen von intelligenten Heuristiken, die helfen, Vertrauenswürdigkeit im Netz zu beurteilen und Fakten von Fakes zu unterscheiden, (ii) die Kenntnis der Geschäftsmodelle von sozialen Medien (wie „pay with your data“) und deren Folgen für die Nutzer, und (iii) das Verstehen der Steuerung des Verhaltens und der Emotionen der Nutzer durch das psychologische Design von Plattformen. Diese kognitiven Kompetenzen benötigen aber auch den Willen, die Fernbedienung für die Steuerung des eigenen Verhaltens wieder selbst in die Hand zu nehmen. Wenn Häuser, Autos und Städte smart werden, warum dann nicht auch die Menschen?

Keynote 2: Prof. Dr. Eckhard Klieme, Preisträger des GEBF-Preises für das Lebenswerk

Zeit: Dienstag, 28.02.2023: 17:30 - 18:30 · *Ort:* R14 R00 A04

Unterrichtsforschung diesseits von Qualitätsdimensionen

Eckhard Klieme
Leibniz Institut für Bildungsforschung und Bildungsinformation, Deutschland

Die heutige Unterrichtsforschung hat ihre Wurzeln in Arbeiten der 1960er und 1970er Jahre, in denen methodisch-didaktische Interventionen erprobt und Interaktionsprozesse erforscht wurden - teils auf der Suche nach allgemeinen Wirkungsfaktoren, teils unter detaillierter Rekonstruktion von sozialen, kognitiven und affektiven Prozessen. Der Bezug zu Allgemeiner und Fach-Didaktik war sichtbar, denn es ging um konkrete Lehr-Lern-Arrangements, um Inhalte, um Verstehen und Missverstehen. Beispielsweise 1974-1980 im ersten DFG-Schwerpunktprogramm unseres Forschungsfeldes.

Die aktuelle Unterrichtsforschung beschäftigt sich intensiv mit generischen Basisdimensionen der Unterrichtsqualität, wie sie um 2000 am MPI für Bildungsforschung formuliert wurden. Damit war ein allgemeiner theoretischer und operationaler Rahmen gewonnen, aber inzwischen gilt die Anschlussfähigkeit zur Didaktik als problematisch - etwa in der Frage, was „Kognitive Aktivierung“ mit unterrichtlichen Praktiken und fachlichen Inhalten zu tun hat. Tatsächlich wurde diese Dimension am MPI als Ergänzung pädagogisch-psychologischer Faktoren (Classroom Management, unterstützendes Klima) um „specific didactic aspects“ diskutiert: „cognitive activation refers to teaching strategies, task settings, and patterns of interaction which allow students to actively construct and cross-link knowledge" (Tätigkeitsbericht 1998-2000).

Der Vortrag wird nach zwei „Schnappschüssen“ aus 1979 und 1999 die Frage stellen: Wie lassen sich in der aktuellen Forschung Qualitätsmessung, Didaktik und Untersuchung der Wissenskonstruktion verknüpfen? Dazu werden Analysen vorgestellt, die sich sehr konkret mit Unterrichtsinhalten, mit spezifischen Unterrichtsmethoden und mit Lerndynamiken innerhalb einer Unterrichtseinheit befassen - insbesondere neue Auswertungen der internationalen TALIS-Video-Studie 2018.

Keynote 3: Prof. Dr. Monika Waldis Weber

Zeit: Mittwoch, 01.03.2023: 9:00 - 10:00 · *Ort:* R14 R00 A04

Gesellschaftswissenschaftliches Lernen in Zeiten der Krise – Wissensordnungen und Kompetenzen revisited

Monika Waldis
Pädagogische Hochschule FHNW, Schweiz

Demokratie steht weltweit unter Druck. Auch in westlichen Gesellschaften sind zuweilen rationale Entscheidungsprozesse gefährdet oder stossen auf mangelnde Akzeptanz. Pandemie, Globalisierung, geopolitische Machtansprüche, Digitalisierung, weitreichende menschliche Eingriffe in das Ökosystem und damit verbundene ökonomische und gesellschaftliche Entwicklungen verstärken Krisenwahrnehmungen und akzentuieren die Notwendigkeit gesellschaftlicher Transformation. In der Folge werden oftmals auch Bildungsmassnahmen vorgeschlagen. So formulieren beispielsweise internationale Organisation wie die UNESCO oder der Europarat mit Bildungskonzepten wie «global citizenship education», «digital citizenship education» oder dem Referenzrahmen für Demokratiekompetenz Erwartungen an die Politische Bildung heranwachsender Bürger:innen. Diese sollen ‘Werte, Einstellungen, Fähigkeiten, Wissen und kritisches Denken’ erwerben, welche zur politischen Partizipation und gesellschaftlichen Teilhabe befähigen. Noch weitgehend offen ist die Integration dieser Konzepte in Bildungsziele der gesellschaftswissenschaftlichen Fächer wie Geschichte und Politische Bildung, denen im Wesentlichen die Förderung politischer Kompetenzen obliegt. Bisherige Curricula und fachspezifische Kompetenzmodelle betonen kognitive Lernziele und epistemologische Einsichten, ausdifferenziert als Frage-, Methoden, Urteils- und Handlungskompetenz. Zu deren Förderung wurden in den letzten Jahren unter anderem Aufgaben zum schriftlichen und mündlichen Argumentieren entwickelt und erprobt. Sie führen Schüler:innen in Perspektivenwahrnehmung, Prozesse der Wissensproduktion, kritisches Denken und deliberative Prozesse des Aushandelns ein. Wird das Anliegen einer Politischen Bildung im Sinne einer

«citizenship education» somit bereits umgesetzt? Müssen die disziplinären Bildungsziele erweitert werden? Welche Bedeutung kommt dabei dem fächerübergreifenden Lernen zu? Im Referat wird evidenzbasiert die Frage erörtert, wie gesellschaftswissenschaftliche Bildung in einer Zeit der Unsicherheit und des raschen gesellschaftlichen Wandels gestaltet und zukünftig ausgerichtet werden soll.

Keynote 4: Prof. Dr. Julika Griem

Zeit: Donnerstag, 02.03.2023: 9:00 - 10:00 · *Ort:* R14 R00 A04

Orte und Räume der Wissenschaftskommunikation

Julika Griem
Kulturwissenschaftliches Institut Essen (KWI), Deutschland

Es ist viel die Rede davon, dass Wissenschaftskommunikation Räume eröffnen und Grenzen und Barrieren überwinden soll; dass sie Menschen abholen und mitnehmen müsse. Von wo aber, und wohin, und auf welchen Wegen? In meinem Vortrag gehe ich davon aus, dass es nicht ausreicht, in verräumlichenden Metaphern über Wissenschaftskommunikation zu sprechen, wenn man sie weiterentwickeln möchte. Ich frage daher konkreter, welche materiellen Orte und Infrastrukturen, welche Bühnen und welche Positionen im Wissenschaftssystem der Kommunikation zur Verfügung stehen bzw. geschaffen werden müssen. Denn auch digitale Wissenschaftskommunikation operiert nicht ortlos: Sie reformatiert existierende Diskursräume wie z.B. das Seminar und die Vorlesung; sie spricht ihr Publikum an konkreten Orten an und vermittelt nicht zuletzt Bilder davon, wie wir heute den Raum und Ort der Universität denken und gestalten wollen. Der Vortrag wird sich anhand verschiedener Beispiele damit befassen, wie Räume und Orte der Wissenschaftskommunikation nicht nur als neutrale Hintergründe und Container fungieren, sondern wie sie Dialog und Disput, das Verstehen und Verhandeln von Wissenschaften in offenen und geschlossenen Konstellationen ermöglichen.

Keynote Nachwuchstagung: Prof. Dr. Matthias Brand

Zeit: Montag, 27.02.2023: 17:00 - 18:30 · *Ort:* S04 T01 A01

Relevante Themen im bildungswissenschaftlichen Kontext: Computerspiele, soziale Netzwerke und mehr - Wann wird aus einer leidenschaftlichen Nutzung des Internets ein Problem?

Matthias Brand
Universität Duisburg-Essen, Deutschland

Die Gaming Disorder wurde unlängst von der WHO als Verhaltenssucht in das internationale Klassifikationssystem von Erkrankungen aufgenommen. Über (Online-) Spiele hinausgehend werden auch andere Internetapplikationen suchtartig genutzt, wie beispielsweise soziale Netzwerkseiten, Online-Pornografie und Online-Shoppingseiten. Der Vortrag beleuchtet psychologische und neurobiologische Mechanismen von Internetnutzungsstörungen.

Session A

Einzelbeiträge

A-01: Krisen als Unterrichtsinhalte

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S06 S00 B29
Chair der Sitzung: Lea Fobel

Politische Maßnahmen in Krisenzeiten – Eine Typologisierung der jungen Bevölkerung im Kontext des Kriegs in der Ukraine

Nina Kolleck, Lea Fobel, Johannes Schuster
Universität Leipzig, Deutschland

Demokratische Einstellungen und Verhaltensweisen sind unverzichtbar für die Erhaltung von Demokratien. Entsprechend bedeutsam sind die Identitäten, Einstellungen und Dispositionen junger Menschen in Bezug auf politische Maßnahmen und Positionierungen (Abs et al. 2017). Jugendliche sind von vielen zentralen Formen der politischen Teilhabe ausgeschlossen, die eine Beachtung junger politischer Auffassungen fördern würde. Doch politische Einstellungen entwickeln sich vor allem in der Phase der Jugend – geprägt durch die Erfahrungen Einzelner in ihren sozialen und politischen Kontexten (Eckstein 2019, S. 417). Die aktuelle junge Generation sieht sich dabei in besonderem Ausmaß geprägt von Krisen – mit der Corona-Pandemie 2020 und dem Krieg in Europa im Frühjahr 2022 als aktuelle Höhepunkte. Studien deuten darauf hin, dass die durch Krisen und Konflikte ausgelöste Angst eine Veränderung von Einstellungen und der öffentlichen Meinung bewirken kann (Schoen 2006). Dabei scheinen sich bestimmte Gruppen herauszubilden, anhand derer sich die jungen Leute typologisieren lassen. Eine Analyse dieser Typen ist von großer Bedeutung, da die Einstellungen der jungen Generation eine wesentliche Rolle für die Zukunft des gesellschaftlichen und politischen Zusammenhalts spielen.

Vor diesem Hintergrund untersucht der Beitrag folgende Fragen: Welche Typen junger Menschen lassen sich hinsichtlich der Bewertung politischer Reaktionen auf den Krieg in der Ukraine bilden? Wie setzen sich diese Typen entlang soziodemografischer Merkmale zusammen?

Im Rahmen der Studie wurde eine repräsentative Online-Befragung mit einem Online-Access-Panel durchgeführt. Es wurden 3240 Personen im Alter zwischen 16 und 29 Jahren im Zeitraum 24.06. bis 26.07.2022 quotenrepräsentativ in allen Bundesländern befragt. Der Fragebogen umfasste u.a. Einstellungen in Bezug auf spezifische Aussagen zu politischen Maßnahmen im Krieg in der Ukraine (bspw. gemeinsame Sanktionen der EU gegen Russland, eine militärische Beteiligung Deutschlands an Kämpfen gegen Russland oder das vollständigen Heraushalten aus dem Konflikt).

Zunächst werden für diesen Beitrag zwei Clusteranalysen (Ward, k-means) durchgeführt, um die Daten bestmöglich zu typologisieren. Anschließend werden die Cluster deskriptiv anhand verschiedener soziodemographischer und politischer Merkmale beschrieben. Abschließend wird eine multinominale Regressionsanalyse durchgeführt, um die sozio-politische Zusammensetzung der Cluster inferenzstatistisch zu bestätigen.

Die Ergebnisse zeigen vier Cluster junger Menschen hinsichtlich ihrer Einstellungen zu politischen Maßnahmen im Krieg in der Ukraine: (1) *Die Diplomatischen*, die besonders EU-einheitliche Sanktionen sowie die Aufnahme der Ukraine in die EU und die Beendigung wirtschaftlicher Beziehungen zu Russland befürworten; (2) *die Allrounder*, die grundsätzlich allen Maßnahmen

zustimmen, sich jedoch in jedem Fall nicht aus dem Konflikt komplett heraushalten wollen; (3) *die Militanten*, die eine Wiedereinführung des Wehrdienstes und eine aktive Beteiligung Deutschlands an den Kämpfen in der Ukraine befürworten sowie radikale Maßnahmen wie eine Niederlegung aller Telefonate mit dem russischen Präsidenten oder Einreiseverbote für russische Bürger:innen nach Deutschland fordern; (4) *die Nationalen*, die sich eher aus dem Konflikt heraushalten wollen und einen größeren Wert auf die Besinnung auf eine deutsche Identität und den Schutz der eigenen Grenzen legen.

Von Cluster 1 bis Cluster 4 nehmen die Zufriedenheit mit der finanziellen Situation sowie die Zufriedenheit mit der Demokratie in Deutschland und das Vertrauen in politische Institutionen konstant ab. Während die *Diplomatischen* vergleichsweise hoch gebildet, jünger und besser situiert sind, zeigen die *Allrounder* tendenziell gegenteilige Ausprägungen. Die *Militanten* sind eher von einem Desinteresse an der Politik geprägt und die *Nationalen* zeichnen sich durch ein geringes Vertrauen in politische Instanzen und eine Unzufriedenheit mit den demokratischen Strukturen Deutschlands aus.

Die multinominale Regressionsanalyse zeigt, dass Bildung, Geschlecht, politisches Vertrauen und die Zufriedenheit mit der Demokratie die Zugehörigkeit zu den jeweiligen Clustern signifikant vorhersagen. Andere Variablen spielen hingegen nur für bestimmte Gruppen eine Rolle. Insgesamt illustrieren die Ergebnisse unserer Studie, dass eine Erfassung der Einstellungen junger Menschen wichtig ist, um die verschiedenen Zielgruppen mit Maßnahmen in der formalen und non-formalen Bildung zu erreichen und Verschwörungsideologien entgegenzuwirken.

„Beteiligt euch!“: Wie evidenzbasierte Informationen und Empfehlungen aus der Klimaforschung den Diskurs über den Klimaschutz beeinflussen

Laura Bilfinger, Benjamin Brummernhenrich, Regina Jucks
Westfällische Wilhelms-Universität Münster, Germany

Obwohl die Evidenz aus der Klimaforschung wenig empirische Unsicherheiten aufweist, werden daraus abgeleitete Empfehlungen nicht grundlegend umgesetzt. Trotz des hohen Stellenwertes von Klimaschutz in der europäischen Bevölkerung fehlt ein konstruktiver Diskurs in der Bildung sowie in der Öffentlichkeit (Eurobarometer, 2020). Diese Studie untersucht, wie das Kommunizieren von Klimaforschung und daraus abgeleiteten Klimaschutzempfehlungen an Individuen und Politik auf Widerstand stoßen kann, welcher wiederum den öffentlichen Diskurs hindern kann. Widerstand gegen sozialen Einfluss lässt sich beispielsweise mit Reaktanz messen, einem motivationalen-psychologischen Zustand, der auftritt, wenn die eigenen Denk- oder Handlungsfähigkeiten bedroht sind und, der dazu führen kann, genau das zu tun, was unerwünscht ist (Brehm, 1966). Erhöhte Reaktanzwerte könnten den konstruktiven Diskurs über den Klimaschutz negativ beeinflussen.

In einem 2x2-Experimentaldesign (N = 623) wurden Klimaschutzempfehlungen für den Verkehrssektor in einem Online-Artikel vor den Hintergrund des neusten Berichtes des Weltklimarats (IPCC, Intergovernmental Panel on Climate Change) präsentiert und von einem Klimawissenschaftler vorgestellt. Erstens wurde variiert, ob der Artikel diese Empfehlungen auf politischer oder individueller Ebene ansprach. Zweitens wurde durch die Variation von hoch- vs. niedrig-kontrollierender Sprache manipuliert, wie direkt der Artikel die Notwendigkeit zum Handeln betonte (z.B. "muss tun" vs. "könnte

tun"). Die abhängigen Variablen, die zur Erfassung der Reaktanz verwendet wurden, waren die wahrgenommene Bedrohung der Freiheit (Dillard & Shen, 2005), Suche nach Gegenargumenten (Moyer-Gusé & Nabi, 2010) sowie die Unterstützung der empfohlenen Klimaschutzmaßnahmen des Online-Artikels. Außerdem wurde durch offene Fragen nach der persönlichen Meinung bezüglich effektiver Klimaschutzmaßnahmen im Verkehr und Wahrnehmung aktueller Berichtserstattung über den Klimawandel gefragt.

Entsprechend unserer Erwartungen, zeigen die Ergebnisse der quantitativen Datenanalyse, Haupteffekte der Art der Klimaschutzempfehlung: Versuchspersonen, die den Artikel mit individuellen Maßnahmen zur Eindämmung des Klimawandels lasen, gaben eine höhere Bedrohung der Freiheit ($p < .001$, $\eta_p^2 = 0.18$), ein höheres Maß an Gegenargumenten ($p < .001$, $\eta_p^2 = 0.06$) und eine geringere Unterstützung für die Empfehlungen an ($p = 0.004$, $\eta_p^2 = 0.01$), im Vergleich zu denjenigen, die den Artikel mit politischen Maßnahmen lasen. Weitere Ergebnisse zeigen, dass nach dem Lesen des Artikels mit hoch kontrollierender Sprache eine höhere Bedrohung der Freiheit wahrgenommen wurde, als wenn die Sprache wenig kontrollierend war ($p < 0.001$, $\eta_p^2 = 0.03$). Entgegen unserer Erwartungen, wurden das Maß an Gegenargumenten und die Unterstützung der Klimaschutzempfehlungen von den sprachlichen Merkmalen nicht beeinflusst. Zudem, wurden keine signifikanten Interaktionseffekte zwischen den Faktoren gefunden. Zwei offene Fragen geben Aufschluss über die individuellen Perspektiven zum Klimawandel.

Die Ergebnisse werden aus einer psychologischen und bildungswissenschaftlichen Perspektive im Hinblick darauf diskutiert, wie das Engagement und die Kommunikation von Einzelpersonen über den Klimaschutz gefördert werden können. Es werden Implikationen für den Umgang mit Evidenz aus der Klimaforschung für die (Hoch-)Schulbildung aufgezeigt.

Bildung für nachhaltige Entwicklung im Schulcurriculum. Eine fächerübergreifende Analyse

Igor Birindiba Batista[1], Katrin Hahn-Laudenberg[1], Hermann Josef Abs[2]
[1]Universität Leipzig, Deutschland; [2]Universität Duisburg-Essen

Ziele/Fragestellung: Der Beitrag untersucht wie „Bildung für nachhaltige Entwicklung" (BNE) in den Curricula der Leitfächer der politischen Bildung sowie weiteren affinen Fächern formell verankert wird. Die curriculumsbezogene Erfassung von BNE-relevanten Aspekten dient als Annährungsmaß für die Repräsentation von Nachhaltigkeitsdimensionen in Lerngelegenheiten für den Kompetenzerwerb (vgl. de Haan & Harrenberg 1999; Rieckmann 2011). Hierbei werden inhaltsbezogene (konzeptionelle Bestandteile von BNE), kompetenzorientierte (Art der Leistungserwartung) und lehrplantheoretische (Art der curricularen Verankerung) Perspektiven berücksichtigt. Konkret wird die Frage bearbeitet, in welchem Ausmaß verschiedene Aspekte eines BNE-Leitbilds in den Kernlehrplänen der Sekundarstufe I erkennbar sind.

Theoretische Rahmung: „Nachhaltigkeit" als gesellschaftlich, ökonomisch und politisch übergreifendes Rahmenkonzept berührt diverse Leitprinzipien der internationalen, europäischen und nationalen Politik mit einer direkten Ausstrahlung und zunehmenden Relevanz im Bildungsbereich (Scheunpflug 2020; Holst et al. 2020; Bormann 2022). Im Beitrag greifen die inhaltsbezogene Perspektivierung mit ihrem thematischen Schwerpunkt (Vorage 2019), die outputbezogene

Differenzierung und ihre kompetenzorientierte Einbettung (Rieckmann 2011) sowie die lehrplantheoretische Perspektive mit ihrer politisch-legitimatorischen BNE-Konzeptualisierung (UNESCO 2021) verschiedene Forschungs- und Diskussionsstränge zum BNE-Themenkomplex auf. Dabei wird die Notwendigkeit beachtet, BNE multimethodisch und multiperspektivisch zu erschließen (vgl. Gräsel et al. 2012; Scheunpflug 2019; Edwards et al. 2020; Stein et al. 2020; Artmaier et al. 2021).

Methode: Am Beispiel nordrhein-westfälischer Schulformen der Sekundarstufe I (GY, GS, RS, HS) wurden pro Schulform sieben Lehrpläne sowie sieben weitere schulformübergreifende Fächer/Rahmenvorgaben (z.B. Praktische Philosophie; Verbraucher*innenbildung) anhand des dreigliedrigen Analyserasters mit der inhaltlich strukturierenden qualitativen Inhaltsanalyse nach Kuckartz & Rädiker (2022) untersucht. Das 1.523 Seiten umfassende Dokumentenmaterial wurde zu 80,9 Prozent von zwei Beurteilenden unabhängig voneinander codiert und weist einen Kappa Wert (nach Brennan & Prediger 1981) von 0,82 auf. Unter Nutzung der Analyse-Software MAXQDA konnten 20 deduktiv-induktiv extrahierte Kategorien (mit 3.125 codierten Segmenten) herausgearbeitet werden, die sowohl eine Bestandsaufnahme über die schul- und fachspezifische Verankerung von BNE als auch eine Dokumentation von Steuerungspotenzialen erlauben.

Ergebnisse: Innerhalb der thematisch-inhaltlichen Analyseebene wird deutlich, dass die normative Setzung, unterschiedliche Anteile der Nachhaltigkeitsdimensionen (Politik, Ökonomie, Ökologie, Soziales, Kultur) sollten gleichberechtigt repräsentiert sein, sich in den untersuchten Dokumenten nicht wiederfindet. So wurden nur wenige Bezüge zur Analyse und Beurteilung politischer Problemstellungen sowie zur Steuerungsfähigkeit gesellschaftlicher Institutionen im expliziten BNE-Kontext identifiziert. Die untersuchten Lehrpläne können Nachhaltigkeit als Konzept nur in beschränktem Maße widerspiegeln. Mit Blick auf die zweite kompetenzorientierte Analyseebene „Gestaltungskompetenz“ konnten teilhabebezogenen und motivationalen Kompetenzen des BNE-Leitbilds (z.B. Gerechtigkeits- und Machtfragen) nur rudimentär erfasst werden. Am schwächsten vertreten waren Teilkompetenzen, die am stärksten die Herbeiführung verbindlicher Entscheidungen in strittigen gesellschaftlichen Fragen adressieren (z.B. Partizipation; Kooperation). Die lehrplantheoretische Analyseebene präsentiert BNE überwiegend als fächerübergreifender Perspektivenwechsel, der neben anderen Konzeptionen wie der Verbraucher*innenbildung koexistiert. Es stellt sich die Frage, inwiefern ein Drittel der untersuchten Lehrpläne ohne die Zuweisung eines bestimmten Inhaltsfelds oder einer expliziten Schwerpunktsetzung dem anspruchsvollen Bildungskonzept BNE mit seiner Mehrperspektivität von Problemlösestrategien und Sachverhalten gerecht werden können.

Diskussion: Aus den vorliegenden stark zusammengefassten Ergebnissen ergibt sich für Politik und Forschung die Implikation, die Vielfalt von BNE-Konzeptionen, einschließlich ihrer ethisch-politischen Positionen und Ziele, stärker in Bildungskontexten zu reflektieren. Ferner gilt es gezielter der Frage nachzugehen, inwiefern in gängigen BNE-Konzeptionen ein Bild vorherrscht, wonach die Verantwortung für ökologische Transformationsprozesse in die individuellen Subjekte verlagert wird und so der Blick auf strukturelle sowie politikgestalterische Aspekte bestehender Verhältnisse verdeckt bleibt.

A-02: Feedback und Kollaboration

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S06 S00 B32
Chair der Sitzung: Livia Kuklick

Computerbasiertes Feedback im digitalen Testen: Welche Rolle spielt das visuelle Design der Feedbackbotschaft?

Livia Kuklick, Marlit Lindner
Leibniz-Institut für die Pädagogik der Naturwissenschaften und der Mathematik (IPN), Germany

Viele Studien zeigen, dass computerbasiertes Feedback positive Effekte auf Lernende in verschiedenen Bildungskontexten hat. Insbesondere unmittelbares elaboriertes Feedback könnte Schülerinnen und Schülern dabei helfen, ihre initial falschen Antworten in späteren Lernstandsabfragen zu korrigieren (z.B. Attali & Van der Kleij, 2017; Enders et al., 2021). Damit könnten Testsituationen sinnvoll angereichert werden und zum Lernen beitragen (z.B. Franklin et al., 2021). Eine Darbietung von unmittelbaren Feedbackbotschaften könnte aber auch Nachteile haben, denn die Verarbeitung von fehlerbezogenem Feedback kostet vermutlich zusätzliche Bearbeitungszeit und könnte sich emotional negativ (auf die positiven und negativen Emotionen der Getesteten) auswirken (vgl. z.B. Kuklick & Lindner, 2021). Wir nehmen an, dass die kognitive Verarbeitung von Feedback und die Emotionen der Testbearbeitenden durch die Ergänzung von aufgabenbezogenen, repräsentationalen Bildern (vgl. z.B. Lindner, 2020) und zusätzlichen Emotional Design Elementen (hier: warme Farben und animierter pädagogischer Agent; vgl. z.B. Brom et al., 2018) in Feedbackbotschaften verbessert werden kann.

In dieser computerbasierten Experimentalstudie haben wir elaboriertes Feedback durch zwei visuelle Designelemente ergänzt (Repräsentationales Bild [RP]; emotionales Design [ED]) und die Wirkung der verschiedenen Feedbackbotschaften auf die Testbearbeitenden untersucht. Unsere Annahme war, dass die Wirkung des Feedbacks auf die Emotionen, die Bearbeitungszeit und den Lernzuwachs durch die vorangegangene Richtigkeit der Antwort substanziell moderiert wird und dass visuell angereichertes Feedback im Vergleich zu textbasiertem Feedback den emotionalen Zustand nach dem Erhalt von (negativen) Feedbackbotschaften verbessert, die Bearbeitungszeit des Feedbacks reduziert, und den Lernzuwachs erhöht.

150 Studierende (69% weiblich; M_{Alter}=25.21; SD_{Alter}=5.86) bearbeiteten einen computerbasierten Test mit 12 Geometrie-Aufgaben (offenes Antwortformat) und erhielten dabei in einem 2×2+1 Between-Subject-Design entweder (1) kein Feedback oder unmittelbares elaboriertes Feedback mit verschiedenen Designelementen: (2) Text, (3) Text + RP, (4) Text + ED oder (5) Text + RP + ED. Wir erfassten die Bearbeitungszeit pro Feedbackbotschaft und die Studierenden bewerteten nach jeder Aufgabe vier Items sowohl zu ihren aktuellen positiven als auch zu ihren negativen Emotionen auf einer 5-Punkte-Likert-Skala. Zur Erfassung des Lernzuwachses diente ein randomisierter Recall-Posttest ohne Feedback. Für die Datenauswertung verwendeten wir moderierte lineare Mixed-Effects Modelle. Eine Post-hoc-Power-Simulation für unsere Mixed-Effects-Analysen (s. Green & McLeod, 2016) ergab eine geschätzte Power von >99%.

Erwartungskonform moderierte die Antwortrichtigkeit auf Ebene der einzelnen Testaufgaben die Wirkung des Feedbacks auf die Emotionen der Getesteten (niedrigeres Maß an positiven Emotionen und höheres Maß an negativen Emotionen nach falschen Antworten, p_s<.001), die Betrachtungszeit

des Feedbacks (längere Verarbeitung nach falschen Antworten, $p<.027$) und die Posttest-Leistung (stärkerer Lernzuwachs nach initial falschen Antworten, $p<.001$). Im Einklang mit vorherigen Studienergebnissen (z.B. Attali & Van der Kleij, 2017; Kuklick & Lindner, 2021) zeigt sich auch in dieser Studie, dass sowohl die kognitiven als auch die affektiven Effekte von Feedback leistungsabhängig sind. Konfirmatorisches Feedback scheint sich affektiv positiver auszuwirken und schneller verarbeitet zu werden als korrektives Feedback.

Hinsichtlich der Designelemente zeigte sich erwartungskonform, dass die Studierenden weniger Zeit auf Feedbackbotschaften mit einem RP ($p=.038$), einem ED ($p=.042$) oder beiden Elementen ($p=.018$) verbrachten als auf rein textbasiertem Feedback, allerdings nur nach falschen Antworten. Möglicherweise halfen die Designelemente also insbesondere bei der Verarbeitung *neuer, korrektiver* Informationen. Keines der visuellen Designelemente beeinflusste jedoch die Emotionen der Studierenden im Vergleich zum Textfeedback ($p_s \geq .381$). Die Feedback-Gruppen unterschieden sich auch nicht signifikant hinsichtlich ihrer Leistung im Posttest ($p_s \geq .381$). Insgesamt scheint visuell angereichertes Feedback affektiv reinem textbasiertem Feedback nicht unbedingt überlegen zu sein. Die geringere Bearbeitungszeit der korrektiven Feedbackbotschaften in Kombination mit dem vergleichbaren Leistungszuwachs im Posttest deutet jedoch darauf hin, dass der Einsatz eines repräsentationalen Bildes und/oder emotionaler Designelementen die Verarbeitungseffizienz von negativen Feedbackbotschaften erhöhen kann. Eine Anreicherung von Feedbackbotschaften mit visuellen Designmerkmalen könnte im Kontext digitaler Leistungstests also vor allem zur kognitiven Unterstützung bei komplexeren Fehlermitteilungen eingesetzt werden.

Automated Feedback on Interest Development: Mediated by Student Perception of Feedback Usefulness

Jan Luca Bahr, Lars Höft, Thorben Jansen
Leibniz Institute for Science and Mathematics Education (IPN), Deutschland

The development of interest in challenging tasks, such as writing an argumentation, is one of the central goals in K-12 science education (Kultusministerkonferenz, 2020a, 2020b, 2020c). Feedback has been shown to trigger students' interest (Harks et al., 2014; Rakoczy et al., 2013; Rakoczy et al., 2019). Adding to this, Koenka et al. (2021) found specific feedback to better foster interest than providing global (i.e., general) feedback. However, students receive less feedback than desirable on written argumentations (Britt et al., 2004) and oftentimes the provided feedback is global. This might make challenging tasks uninteresting to students. Online learning environments have attempted to stimulate interest through providing adaptive (i.e., matching) and specific (i.e., well elaborated) feedback. However, little is known about the student perspective within the feedback process. That is, do students perceive adaptive or specific feedback as more useful than global feedback? And how does this relate to student's interest development?

We hypothesized that both adaptive (H1) and non-adaptive specific feedback (H2) would be perceived as more useful than global feedback. We expected that perceived usefulness would relate positively to changes in interest after feedback (H3) and after students revised their texts (H4). Correspondingly, we hypothesized that there would be indirect effects on changes in interest after feedback (H3a) and after students revised their texts (H4a) via students' perception of feedback

usefulness. Further, we expected that both adaptive feedback and specific feedback would be related to a positive change in interest after feedback (H5) and after revision (H6).

Methods

We conducted an experimental study with a between-subject design and three measurement points. 463 German students (*Mage* = 17.28, *SD* = 1.03, age range = [14, 21]; 55% female) from higher secondary schools (*Gymnasium*) or comprehensive schools (*Gesamtschule*) from ninth to 13th grade participated in the study.

Procedure

Students wrote an argumentative text about a climate change related topic. Situational interest was assessed after having seen the assignment, after having written a first draft and after revision. Students were given feedback on their first draft. The control group (i.e., global feedback) received feedback stating there was room for improvement without specifically pointing out on where to improve. Whereas the two experimental groups saw one of four feedbacks specifically addressing quality indicators of argumentations. In the adaptive group, an algorithm created from a pilot study was used which selected the best matching feedback message in 80% of cases using a bag-of-words approach (see Wang et al., 2020). In the specific group, the feedbacks were randomly assigned. Afterwards, the students rated the received feedback according to its usefulness.

Results

All reported findings refer to the model shown in Figure 1. We found support for H1, H2, H3, H3a, H4 and H4a. Contrary to expectations, we did not find support for H5 and H6.

Significance of the study

In line with earlier findings from traditional classroom settings (Harks et al., 2014; Rakoczy et al., 2013; Rakoczy et al., 2019) we have shown the impact of perceived feedback usefulness on interest development within online learning environments. Underlining its importance, we found the impact to persist even after revision. Moreover, we demonstrated that specificity of feedback is key to student interest development as its perceived as useful. This is particularly important since we do not have the technical means yet to provide adaptive (i.e., matching) automated feedback in all cases. Unexpectedly, random specific feedback was perceived as more useful than adaptive specific feedback. This will be discussed during the conference when the actual fit of feedback and student text in both experimental conditions is calculated.

Die Rolle eines Kooperationsskripts für die Veränderung kooperationsbezogener Selbstwirksamkeitserwartung in einer computerbasierten Lernumgebung

Joyce Hobrecht, Julia Eberle, Isis Tunnigkeit, Leonie vom Bovert, Sebastian Strauß, Marcel Schmittchen, Arlind Avdullahu, Nikol Rummel
Ruhr-Universität Bochum, Germany

Der Bedarf an globaler Zusammenarbeit zwischen Expert:innen unterschiedlicher Domänen steigt zunehmend (Meier, 2005). Eine zentrale Herausforderung ist dabei der Mangel an adäquaten internalen Skripts über effektive computerunterstützte Kooperation und das entsprechende prozedurale Wissen (Kollar et al., 2007). Ein weiterer, jedoch bislang wenig untersuchter Faktor ist die

kooperationsbezogene Selbstwirksamkeitserwartung. Diese beschreibt den Grad des Vertrauens eines Individuums in die eigene Fähigkeit, in kooperativen Situationen wirksam handeln zu können (Mattson, 2011). In Bezug auf andere Kontexte ist bekannt, dass die Bedeutung von Selbstwirksamkeit für die Bewältigung von Herausforderungen vergleichbar mit der des Vorwissens ist (Bandura, 1997; Mattson, 2011; Pajares, 1996). Für den komplexen Kontext von computerunterstützter interdisziplinärer Kooperation fehlen entsprechende Untersuchungen bisher noch weitgehend. Insbesondere fehlen Erkenntnisse darüber, wie eine realistische Selbstwirksamkeitserwartung im Kooperationskontext gefördert werden kann.

Ein vielversprechender Ansatz zur Beeinflussung der Selbstwirksamkeitserwartung im Kooperationskontext ist der Einsatz von externalen Kooperationsskripts, einer Form des Scaffoldings (Kollar et al., 2006), die der Strukturierung und Anregung wirkungsvoller Kooperationsprozesse dient (King, 2007). Scaffolding zeigte sich bereits hinsichtlich verschiedener Aufgabenfelder als eine wirkungsvolle Strategie zur Förderung der Selbstwirksamkeitserwartung (Yantraprakorn et al., 2013; Grothérus et al., 2019; Prabawanto, 2018; Valencia-Vallejo et al., 2019). Aus diesem Grund werden Kooperationsskripts im Kontext von Kooperationsprozessen als eine vielversprechende Fördermaßnahme für kooperationsbezogene Selbstwirksamkeitserwartung vermutet.

Vor diesem Hintergrund wurden in der vorliegenden Studie drei Fragen untersucht:

1. Inwiefern unterscheidet sich die Veränderung von kooperationsbezogener Selbstwirksamkeitserwartung im Rahmen interdisziplinärer computerbasierten Kooperationsprozesse nach einer Lernphase mit Unterstützung durch ein externales Kooperationsskript?
2. Welche Erfahrungen im Kooperationsprozess erklären die Veränderung der Selbstwirksamkeitserwartung?
3. Inwiefern fördert ein externales Kooperationsskript diese relevanten Erfahrungen im Kooperationsprozess?

Zur Beantwortung der Forschungsfragen wurde ein Teildatensatz einer Online-Laborstudie untersucht. Im Rahmen der Studie bearbeiteten die Studierenden nach einem Pretest in Dreiergruppen ein fiktives Problemszenario (Lernphase), füllten einen Zwischentest aus, bearbeiteten ein zweites Problemszenario (Testphase) und schlossen mit einem Posttest ab. Der Teildatensatz umfasst 72 Studierenden, die in der Lernphase entweder keine Unterstützung erhielten oder durch ein Kooperationsskript unterstützt wurden, welches eine Phasierung des Kooperationsprozesses sowie Prompts zur Herstellung und Aufrechterhaltung effektiven Kooperationsverhaltens umfasste. Die kooperationsbezogene Selbstwirksamkeitserwartung, als zentrale abhängige Variable, wurde durch einen Fragebogen im Pre- und Posttest erfasst. Die Erfahrungen der Proband:innen im Kooperationsprozess wurden anhand eines Kodierschemas erfasst. Zur Auswertung der Daten wurden ANOVAs mit Messwiederholung sowie Regressionsanalysen eingesetzt.

Die Ergebnisse zeigten keinen Unterschied zwischen der Kontrollbedingung und der Bedingung mit Kooperationsskript hinsichtlich der Veränderung der kooperationsbezogenen Selbstwirksamkeitserwartung ($F(1, 70) = 1.364$, $p = .247$). Weiterhin zeigte sich, dass ein gehäuftes

Ausführen technischer Koordinationsprozesse während der Zusammenarbeit einen kritischen Faktor für das Absinken kooperationsbezogener Selbstwirksamkeitserwartung darstellt ($F(2,71)=4.14$, $p=.020$). Diese technische Koordination trat in der Kooperationsskript-Bedingung signifikant häufiger auf als in der Kontrollbedingung ($t(52,224) = 1.987$, $p = .026$).

Explorative Einzelfallanalysen von Proband:innen, die einen besonders hohen Anstieg bzw. Abfall der Selbstwirksamkeitserwartung aufwiesen zeigten, dass besonders dann starke positive bzw. negative Veränderungen in der Selbstwirksamkeitserwartung auftraten, wenn die eigene Selbstwirksamkeit vor der kooperativen Problemlösung überdurchschnittlich niedrig bzw. hoch eingeschätzt wurde. Ferner zeigte sich, dass Proband:innen mit hohem Anstieg der Selbstwirksamkeitserwartung die Kooperationsaufgabe als kognitiv belastender und anspruchsvoller wahrnahmen, mit dem Ergebnis des Kooperationsprozesses jedoch deutlich zufriedener waren, als die Proband:innen mit starkem Absinken der kooperationsbezogenen Selbstwirksamkeitserwartung.

Zusammenfassend zeigen die Ergebnisse, dass das eingesetzte Kooperationsskript keine positiven Auswirkungen auf die Selbstwirksamkeitserwartung hatte. Vielmehr führte das Skript durch seine technische Komplexität zu erhöhtem technischen Koordinationsbedarf in der Zusammenarbeit. Diese Ergebnisse sensibilisieren für mögliche Gefahren beim Einsatz externaler Kooperationsskripts im Kooperationsprozess. Die explorativen Einzelfallanalysen deuten zudem darauf hin, dass während der Kooperation ein Abgleich zwischen den Einschätzungen der eigenen Wirksamkeit und den tatsächlichen Kooperationskompetenzen stattgefunden hat und dass die wahrgenommene Komplexität der Kooperationsaufgabe einen veränderungsrelevanten Faktor für kooperationsbezogene Selbstwirksamkeitserwartung darstellen kann.

Ist diese Webseite wirklich verlässlich? Effekte eines Kollaborationsskripts und eines Reflexionsprompts auf die Evaluation von Internetquellen

Carolin Baumgarten[1], Heiko Krabbe[1], Ingo Kollar[2], Marc Stadtler[1]
[1]Ruhr-Universität Bochum, Deutschland; [2]Universität Augsburg, Deutschland

Theoretischer Hintergrund

Heranwachsende nutzen zunehmend Internetquellen für Bildungszwecke (Eickelmann et al., 2019). Die Evaluation von Internetquellen fällt ihnen jedoch häufig schwer und bedarf daher der Förderung (Braasch et al., 2013). Obwohl kooperative Lernformen das Potenzial besitzen, eine besonders tiefe Elaboration von Lerninhalten anzuregen, haben diese in der bisherigen Forschung zur Förderung von Quellenbewertungskompetenzen wenig Beachtung gefunden. In der vorliegenden Studie adressieren wir diese Forschungslücke, indem junge Erwachsene kooperativ Rechercheaufgaben im Internet bearbeiteten. Dabei variieren wir, ob die Lernpartner:innen zur Strukturierung ihrer Zusammenarbeit ein externes Skript erhalten (Mayweg-Paus et al., 2020). Zudem variieren wir, ob die Lernpartner:innen zur Reflexion der Interaktion aufgerufen werden, um lernförderliche Kollaborationsformen auch nach Ausschleichen des Skripts aufrecht zu erhalten (Fischer et al., 2013).

Wir erwarten förderliche Effekte des Kollaborationsskripts auf die Kommunikation zwischen den Lernpartner:innen und auf die resultierenden individuellen Kompetenzen der Quellenbewertung (u.a. Wissen über Strategien der Quellenbewertung, Verlässlichkeitsurteile und

Selbstwirksamkeitserwartung hinsichtlich der Quellenbewertung). Diese Effekte sollten besonders ausgeprägt sein, wenn die Skriptinternalisierung durch einen Reflexionsprompt angeregt wird.

Methode

Stichprobe und Design

An der Studie nahmen 118 Universitätsstudierende (*M*=23,2 Jahre; *SD*=3,2; 31%m, 68%w, 1%d) mit geringem Vorwissen zu den Recherchethemen teil. Die Proband:innen wurden randomisiert einer von vier Versuchsbedingungen zugewiesen (mit/ohne Kollaborationsskript; mit/ohne Reflexionsprompt).

Materialien

Rechercheinstruktion und Material. Alle Dyaden bearbeiteten in zwei aufeinanderfolgenden Phasen Rechercheaufgaben zu zwei kontroversen Wissenschaftsthemen. Zu jedem Thema erhielten sie vier vorausgewählte authentische Webseiten (je zwei verlässlich/nicht verlässlich). In der Rechercheinstruktion wurden die Dyaden über Kriterien aufgeklärt, die eine verlässliche Quelle ausmachen und aufgefordert, auf den vorgegebenen Webseiten verlässliche Informationen zu identifizieren.

Kollaborationsskript. Die Hälfte der Dyaden wurde in der Übungsphase durch ein Kollaborationsskript zum argumentativen Diskurs über Quelleninformationen angeleitet. Das Skript fordert die Dyadenpartner:innen zunächst auf, sich einzeln mit je zwei Webseiten auseinanderzusetzen und diese entlang von vier Leitfragen zu bewerten, die die zuvor eingeführten Verlässlichkeitskriterien aufgreifen. In anschließenden Partnerinterviews geben die Proband:innen über ihre Quellenbewertungen Auskunft, zeigen ihre Vorgehensweise und verständigen sich auf ein gemeinsames Verlässlichkeitsurteil für jede der vier Webseiten.

Reflexionsprompt. Am Ende der Übungsphase erhielt die Hälfte der Proband:innen eine Aufforderung zur Reflexion ihrer Zusammenarbeit. Der Prompt fordert dazu auf, als hilfreich empfundene Faktoren der Zusammenarbeit zu benennen und festzuhalten, welche dieser Elemente in der anschließenden Rechercheaufgabe (Testphase) beibehalten werden sollen.

Versuchsablauf und Messinstrumente

Die Studie wurde über ein Videokonferenztool durchgeführt. Die Messungen erfolgten computerbasiert in einem Prä- und Posttestdesign. Im Prätest wurden Kontrollvariablen wie das selbsteingeschätzte Vorwissen und die Selbstwirksamkeitserwartung hinsichtlich der Quellenbewertung erfasst. Danach fand die 45-minütige Übungsphase statt, in der die Dyaden zum ersten Recherchethema arbeiteten und die experimentellen Manipulationen umgesetzt wurden. In der darauffolgenden 45-minütigen Testphase recherchierten alle Dyaden ohne Unterstützungsmaßnahmen zu einem neuen Thema (Testphase). Die aufgezeichneten Dyadeninteraktionen der Testphase wurden hinsichtlich der Häufigkeit der Quellenevaluationen inhaltsanalysiert (Cohen’s kappa >.81). Der Posttest, den die Proband:innen individuell bearbeiteten, diente zur Erfassung der abhängigen Variablen (u.a. Wissen über Strategien der Quellenbewertung, Verlässlichkeitsurteile und Selbstwirksamkeitserwartung hinsichtlich der Quellenbewertung).

Ergebnisse

Wie erwartet, wiesen die Dyadeninteraktionen derjenigen, die mit Kollaborationsskript und Reflexionsprompt arbeiteten, die meisten Quellenevaluationen auf, $F(1,112)=10,05, p=.002, part\eta^2=.082$. Zudem konnten Proband:innen der Skriptbedingung mehr Quellenevaluationsstrategien erinnern als Proband:innen, die ohne Skript arbeiteten, $F(1,114)=20,86, p=<.001, part\eta^2=.15$. Proband:innen, die mit Reflexionsprompt arbeiteten, erinnerten entgegen der Erwartung weniger Evaluationsstrategien als diejenigen, die ohne Reflexionsprompt arbeiteten, $F(1,114)=4,26, p=.041, part\eta^2=.036$. Die Analyse der Verlässlichkeitsratings der Webseiten zeigte keine Bedingungsunterschiede, alle $F<2,43, p>.12$. Die Selbstwirksamkeitserwartung stieg über alle Bedingungen hinweg an $F(1,114)=11,55$, $p=<.001, part\eta^2=.09$.

Diskussion

Die Ergebnisse zeigen, dass das Wissen über Strategien der Quellenbewertung sowie deren Nutzung im dyadischen Austausch mithilfe des Kollaborationsskripts gefördert wird. Die Verlässlichkeitsratings zeigen jedoch, dass weder das Skript noch der Reflexionsprompt die Quellenbewertung signifikant fördern konnte. So bedarf es weiterführender Forschung hinsichtlich der Gestaltung eines wirksamen Kollaborationsskripts im Kontext der Quellenbewertung.

A-03: Struktur Bildungssystem

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S06 S00 B41
Chair der Sitzung: Daniel Deimel

Verbindung von Landesstatistik mit Schulleistungsstudien: Sozialräumliche Effekte auf die intendierte politische Partizipation von Jugendlichen

Daniel Deimel, Hermann J. Abs
Universität Duisburg-Essen, Deutschland

Hintergrund

Politische Partizipation bezeichnet alle individuellen und gemeinschaftlichen Handlungen von Bürger:innen mit dem Ziel, direkt und indirekt politische Entscheidungen zu beeinflussen (Theocharis & van Deth, 2018). Die Hinwirkung auf die Bereitschaft und die Befähigung zur politischen Partizipation ist als Ziel im Bildungswesen in demokratischen Gesellschaften verankert (Hoskins, 2006). Gleichzeitig entwickeln junge Menschen diese Bereitschaft in unterschiedlichem Ausmaß. Die Frage ist, welche Kontexte hierfür verantwortlich sind. Schulische Kontexte des Aufbaus von politischer Partizipationsbereitschaft sind gut beschrieben (Dassonneville et al., 2012; Geboers et al., 2013). Weniger Klarheit besteht hingegen über die Rolle des Sozialraumes, in dem junge Menschen sozialisiert werden.

Sozialräumliche Charakteristika, wie die soziale Komposition, stehen in Zusammenhang mit schulischer Leistung (Jeworutzki & Schräpler, 2019). Auch für die Entstehung von politischer Partizipationsbereitschaft lassen sich Effekte des Sozialraumes theoretisch annehmen. Im Allgemeinen stehen strukturelle Merkmale einer Region mit Lebenschancen und (Un-)Zufriedenheit in Zusammenhang, die sich in bestimmten Partizipationsmustern wie einer Protestwahl äußern können (Holtmann, 2019). Außerdem unterscheiden sich städtische und ländliche Räume hinsichtlich des sozialen Kapitals dieser Gemeinschaften (Putnam, 2000), das wiederum einen Mobilisierungseffekt für politische Partizipation haben kann (Verba et al., 1995). Nicht-traditionelle Lebensstile und somit die Einbindung etwa in kirchliche Mobilisierungsnetzwerke sind abseits von Städten weniger verbreitet (Spellerberg, 2014). Auf der anderen Seite bieten urbane Räume möglicherweise mehr Gelegenheiten zu weniger konventionellen politischen Aktivitäten, wie die Möglichkeit zur Teilnahme an Protesten (Hasson, 1997; Schoene, 2018).

Analog hierzu schließt die International Civic and Citizenship Education Study (ICCS 2009 & 2016; Schulz et al., 2010; Schulz et al., 2018) sozialräumliche Erklärungsmuster für Unterschiede in den Ergebnissen politischer Sozialisationsprozesse ein, die durch das unterschiedliche Vorhandensein von ökonomischen, kulturellen und sozialen Ressourcen plausibel werden. Empirische Hinweise zu dieser multidimensionalen Betrachtung des Sozialraumes bleiben jedoch aus, lediglich ein Einfluss der Stadtgröße konnte hinsichtlich der Ausprägung politischer Kompetenz (Coopmans et al., 2020) und der Wahlbereitschaft (Isac et al., 2014) beschrieben werden.

Methoden & Ergebnisse

Dieser Lückenschluss gelingt durch das Design dieser Arbeit am Beispiel des Bundeslandes Nordrhein-Westfalen. Es werden regionalstatistische Daten zu verschiedenen sozio-ökonomischen, demografischen und ökonomischen Themen auf der NUTS-3 Ebene, also auf der Ebene der Stadt oder des Landkreises (IT.NRW, 2020a, 2020b) mit dem nordrhein-westfälischen Subsample der ICCS

2016 (Abs & Hahn-Laudenberg, 2017) verbunden. Die Stichprobe besteht aus einem repräsentativen Sample von 1.451 Schüler:innen in der achten Jahrgangsstufe in 59 Schulen. Zunächst wird eine mehrdimensionale Beschreibung des Sozialraumes vorgenommen. Hierzu werden insgesamt 25 Indices gebildet, die nach einer Hauptkomponentenanalyse drei Sozialindikatoren zugeordnet werden: dem Grad der Verstädterung, dem Grad des aggregierten individuellen Wohlstandes sowie der Vielfalt des Chancenangebots (siehe Abb. 1). In einem zweiten Schritt werden Mehrebenenmodelle definiert, welche unter Kontrolle von mittlerem SES und Schulform Effekte dieser Sozialindikatoren auf wahlbezogene sowie zivil ungehorsame Partizipationsbereitschaft vorhersagen sollen (Deimel & Abs, 2022).

Im Ergebnis zeigt sich, dass vor allem die Vielfalt des Chancenangebots des Sozialraumes einen Effekt hat. Je höher dieses ist, desto seltener äußern Schüler:innen die Bereitschaft, an Wahlen teilnehmen zu wollen, sobald sie erwachsen sind. Gleichzeitig äußern Schüler:innen in diesen Sozialräumen häufiger die Absicht, zukünftig zivil ungehorsame Partizipationsakte ausüben zu wollen. Hinsichtlich des aggregierten individuellen Wohlstandes und dem Grad der Urbanisierung zeigt sich hingegen kein Effekt, wobei mögliche Effekte der sozioökonomischen Zusammensetzung durch die Kontrollvariablen abgedeckt wurden. Eine Limitierung des gewählten Ansatzes zeigt sich darin, dass die Auflösung der genutzten regionalstatistischen Daten auf der NUTS-3 Ebene vergleichsweise grob ist. Somit zeigt sich der analysierte sozialräumliche Einfluss eher mittelbar in lokal unterschiedlichen Sozialisationsbedingungen, deren Wirkzusammenhänge einer vertieften Betrachtung bedürften. Dennoch wird ersichtlich, dass der Ansatz, verschiedene Datenquellen zur Erklärung von Unterschieden von Large Scale Assessment Daten vielversprechend ist, vor allem, wenn hinsichtlich der Kontexte jenseits der schulischen Bedingungen nur begrenzte Daten zur Verfügung stehen.

Studienentscheidung, soziale Herkunft und institutioneller Habitus – Wie Schulkulturen den Übergang in die Hochschule beeinflussen

Annabell Daniel[1,3], Frederick de Moll[2], Victoria Zeddies[3], Jan Scharf[3], Nicolas Hübner[4], Fabian Wolff[5], Martin Daumiller[6]
[1]Ludwig-Maximilians-Universität München, Germany; [2]Universität Bielefeld; [3]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; [4]Universität Tübingen; [5]Universität Koblenz-Landau; [6]Universität Augsburg

Theoretischer Hintergrund

Die Aufnahme eines Hochschulstudiums wird sowohl aus theoretischer als auch empirischer Perspektive zumeist als das Ergebnis rationaler Entscheidungskalküle betrachtet (Boudon 1974; Watermann et al. 2014), während der schulische Kontext, innerhalb dessen junge Erwachsene ihre Studienentscheidung treffen, deutlich seltener Beachtung findet. Internationale Befunde zeigen allerdings, dass der Anteil der Schüler:innen, die nach der Schule die Aufnahme eines Studiums planen, auch nach Kontrolle individueller Merkmale wie Herkunft und Leistung erheblich zwischen Schulen variiert (Pustjens et al. 2004; Robinson & Roksa 2016). Eine mögliche Erklärung hierfür bietet das Konzept des institutionellen Habitus (Bourdieu 1984), demzufolge Schulen gewisse Werte und Normen transportieren, die ein Hochschulstudium für einzelne Schüler:innen sehr wahrscheinlich, denkbar oder gar unmöglich erscheinen lassen (Reay et al. 2001; Smyth & Banks 2012). Qualitative Studien verweisen zudem auf verschiedene Dimensionen einer Kultur der Studienorientierung, wie studienvorbereitende Maßnahmen oder auch Einstellungen schulischer Akteure, die einen

Hochschulzugang begünstigen können (z.B. McDonough 1997). Inwieweit auch in einem hochgradig stratifizierten Bildungssystem wie Deutschland der schulische Kontext, jenseits von Schulformeffekten (Schuchart 2019) die Studienentscheidung beeinflusst, ist weitgehend offen.

Fragestellungen

Die vorliegende Studie geht daher der Frage nach, inwiefern sich zusätzlich zu Effekten der sozialen und leistungsbezogenen Komposition in der gymnasialen Oberstufe Effekte einer Kultur der Studienorientierung auf den Übergang in die Hochschule identifizieren lassen. Dabei berücksichtigen wir sowohl eine normative Dimension (Studienaspirationen der Mitschüler:innen) als auch eine organisationale Dimension (z.B. Maßnahmen der Studienvorbereitung). Zudem prüfen wir differenzielle Effekte auf die Studienentscheidung in Abhängigkeit der sozialen Herkunft. Angenommen wird, dass eine ausgeprägte Kultur der Studienorientierung an den Schulen insbesondere jungen Erwachsenen mit geringerem sozioökonomischen Status dabei helfen könnte, Informationsdefizite und einen Mangel an sozialem Kapital zu kompensieren (Achinstein et al. 2015) und ihre Entscheidung für ein Studium entsprechend begünstigt.

Methode

Als Datengrundlage dient die Startkohorte 4 des Nationalen Bildungspanels (NEPS; Blossfeld & Roßbach, 2019). Die Längsschnittstichprobe umfasst 3466 Schüler:innen (davon 56% weiblich), die sich auf 152 Schulen mit gymnasialer Oberstufe verteilen. Untersucht wird, inwiefern (a) die aggregierte Studienaspiration der Mitschüler:innen (als Indikator für normative Werte) und (b) durch die Schule angebotene Hochschulbesuche, Studienberatungsgespräche mit Lehrkräften sowie die mittlere studienpropädeutische Kompetenz (als organisationale Maße einer Kultur der Studienorientierung) die Aufnahme eines Hochschulstudiums etwa eineinhalb Jahre nach Schulabschluss beeinflussen. Die Auswertung erfolgte anhand von Mehrebenenmodellen, in denen für die soziale und leistungsbezogene Komposition sowie die individuelle Studienaspiration beim Eintritt in die Oberstufe kontrolliert wird.

Ergebnisse

Die Ergebnisse zeigen, dass insgesamt zwar nur ein geringer Anteil der Varianz in der Studienentscheidung auf Unterschiede zwischen Schulen zurückzuführen ist, aber dass der institutionelle Habitus durchaus relevant für die Wahl nachschulischer Bildungswege ist. Während die soziale Zusammensetzung und die Aspirationen der Mitschüler:innen eher unbedeutend waren, zeigte sich unter Kontrolle der individuellen Leistung ein negativer Effekt der aggregierten Mathematikleistungen auf die Studienentscheidung. Bezogen auf die organisationale Dimension einer Kultur der Studienorientierung erwiesen sich insbesondere Studienberatungsgespräche mit Lehrkräften als bedeutsam. Entgegen unserer Annahmen profitierten von diesem Angebot jedoch vor allem sozial privilegierte Schüler:innen. Damit deuten unsere Befunde darauf hin, dass der institutionelle Habitus zwar Übergänge in die Hochschule begünstigt, aber soziale Ungleichheiten nicht nur reproduziert, sondern sogar verstärken kann (Jæger & Breen 2016).

Schulen als differenzielle Entwicklungsmilieus: Wie schulische Bildungsgänge und Schulkomposition den Übergang nach Ende der Pflichtschulzeit beeinflussen – ein Kohortenvergleich

Katja Scharenberg, Wolfram Rollett
Pädagogische Hochschule Freiburg, Deutschland

Theoretischer Hintergrund

Bildungssysteme unterscheiden sich im Zeitpunkt der externen Leistungsdifferenzierung in unterschiedliche Bildungsgänge, die in den meisten OECD-Ländern nach Ende der Pflichtschulzeit erfolgt (Woessmann, 2009). Ländervergleiche zeigen eine höhere Bildungsungleichheit bei früher Leistungsdifferenzierung (z.B. Hanushek & Woessmann, 2005). Ein Grund hierfür ist, dass externe Leistungsdifferenzierung zu sozialer Segregation und Stratifizierung der Schülerschaft führen kann (Baumert et al., 2006): Je nach besuchtem Schultyp haben Schüler:innen unterschiedliche Entwicklungschancen.

Empirische Studien haben Effekte differenzieller Lern- und Entwicklungsmilieus auf leistungsbezogene Kriteriumsvariablen wiederholt nachgewiesen (z.B. Baumert et al., 2006; Gröhlich et al., 2010; Scharenberg et al., 2014). Demnach kommt es in der Sekundarstufe zur Öffnung der Leistungsschere zwischen verschiedenen Bildungsgängen und zur kumulativen Privilegierung oder Benachteiligung in Abhängigkeit von Schultyp oder Schülerkomposition. Auch für die Schweiz liegen entsprechende Befunde vor (Neumann et al., 2007), während dies für den Übergang nach der Pflichtschulzeit bislang kaum systematisch untersucht wurde. An dieser Übergangsschwelle müssen alle Schüler:innen eine Bildungsentscheidung treffen und sich für ein allgemein- oder berufsbildendes Ausbildungsangebot entscheiden (Jüttler et al., 2021). Doch nicht alle Schüler:innen sind hierbei erfolgreich (Hupka-Brunner & Meyer, 2021; Scharenberg et al., 2016).

Untersuchungskontext unserer Studie ist die Schweiz, wo die Sekundarstufe I je nach kantonalen Gegebenheiten in verschiedenen Strukturmodellen organisiert ist (Schweizerischer Bildungsserver, 2021). Das gestufte Modell mit in allen Fächern nach Leistungsanforderungen getrennten Schultypen ist das häufigste Strukturmodell, das aber wenig Durchlässigkeit ermöglicht (EURYDICE, 2021).

Fragestellung

Es wurden folgende Forschungsfragen untersucht:

1. Übt die Schultypzugehörigkeit einen Effekt auf die Wahrscheinlichkeit eines erfolgreichen Übergangs nach Ende der Pflichtschulzeit aus, der über Effekte individueller Merkmale hinausgeht?
2. Hängen Unterschiede in der Schulzusammensetzung mit der Wahrscheinlichkeit eines erfolgreichen Übergangs zusammen?
3. Hat die Schulzusammensetzung einen Einfluss auf den erwarteten Effekt der Schulformzugehörigkeit auf die Wahrscheinlichkeit eines erfolgreichen Übergangs?

Methode

Datengrundlage ist die Schweizerische Mehrkohorten-Längsschnittstudie „Transitionen von der Erstausbildung ins Erwerbsleben" (TREE). Die erste Kohorte (TREE1) untersuchte Jugendliche, die an PISA 2000 teilnahmen und im selben Jahr aus der obligatorischen Schulpflicht entlassen wurden. Die zweite TREE-Kohorte ist die Schulabgängerkohorte aus dem Jahr 2016. Die Analysestichprobe umfasst n=4707 Schüler:innen in 208 Schulen (TREE1) bzw. n=5239 Schüler:innen in 316 Schulen (TREE2). Die Daten wurden gewichtet (Sacchi, 2013; Sacchi, in Vorb.), um Verzerrungen zu

kompensieren, die sich aus ungleichen Ziehungswahrscheinlichkeiten und selektivem Stichprobenschwund im Längsschnitt ergeben können (Little & Rubin, 2020).

Abhängige Variable ist der erfolgreiche bzw. nicht erfolgreiche Übergang in eine post-obligatorische zertifizierende allgemein- oder berufsbildende Ausbildung im ersten Jahr nach Schulende. Logistische Mehrebenenmodelle (HLM 8.2.2.1; Raudenbush et al., 2019) kontrollieren auf Individualebene das Geschlecht, die Sprache im Haushalt (Kantonssprache ja/nein), den Sozialstatus (HISEI; Ganzeboom et al., 1992) sowie standardisierte Mathematik-Testleistungen (PISA 2000: Adams & Wu, 2002; ÜGK 2016: Angelone & Keller, 2019) sowie die Schultypzugehörigkeit (Grund- vs. erweiterte Anforderungen). Auf Aggregatebene wurden die durchschnittliche leistungsbezogene und soziale Schulzusammensetzung sowie der Anteil der Schüler:innen, die keine offizielle Kantonssprache sprechen, berücksichtigt. Schrittweise Modellspezifikationen wurden mittels Devianzstatistik verglichen (Hox et al., 2018). Fehlende Werte wurden modellbasiert (m=5) geschätzt.

Ergebnisse

Die Varianzzerlegung ergab signifikante Varianzkomponenten (TREE1: ρ=0.217; TREE2: ρ=0.128; jeweils p<.001), die die Notwenigkeit verdeutlichen, hierarchische Datenstrukturen zu berücksichtigen. Auf Individualebene erwies sich unter Kontrolle der anderen Variablen die Schulformzugehörigkeit als bedeutsamer Prädiktor: Schüler:innen an Schulen mit erweiterten Anforderungen hatten höhere Übergangswahrscheinlichkeiten (TREE1: OR=1.87; TREE2: OR=2.96; jeweils p<.001). Auf Schulebene zeigten sich im vollständig spezifizierten Modell u.a. Vorteile bei einer sozial privilegierteren (TREE1: OR=3.21, p<.001) bzw. leistungsstärkeren Schülerschaft (TREE2: OR=1.33, p=.042). Die Kompositionseffekte waren jedoch konfundiert und traten z.T. bildungsgangspezifisch auf.

Diskussion

Unsere Befunde erweitern den bisherigen empirischen Erkenntnisstand, indem sie einen Schlüsselindikator der Bildungsbiographie von Schüler:innen berücksichtigen und unterstreichen die Bedeutung kompositioneller und institutioneller Merkmale für diesen Übergang, der eine der wichtigsten Entwicklungsaufgaben im (jungen) Erwachsenenalter darstellt (Havighurst, 1972).

Die Rolle der Schulaufsicht bei der Bewältigung der Auswirkungen der COVID-19 Pandemie auf Schule und Bildung

Stephan Gerhard Huber[1], Robert Pham-Xuan[2], Bettina Gördel[3], Manuela Egger[1], Nadine Schneider[1], Isabella Lussi[4]
[1]Pädagogische Hochschule Zug, Schweiz; [2]Universität Innsbruck, Österreich; [3]Deutschland; [4]Interface Politikstudien Luzern, Schweiz

Einleitung und theoretischer Hintergrund: Die COVID-19 hat im Bildungssystem schnelle Handlungen erfordert. Es herrschte grosse Unsicherheit aufgrund dieser ungewissen Lage und Entwicklung der Pandemie. Dennoch musste gehandelt werden. Entscheidungsträger*innen der Schulaufsichten waren stark im Krisenmanagement gefordert.

Wie frühere Studien zeigen (z.B. Johansen et al., 2012; Pearson et al., 2007; Miller, 2006; Rosenthal & Kouzmin, 1993), sind Planung, Aufgabenverteilung und Commitment wesentliche Bestandteile zur Bewältigung gängiger Auswirkungen einer Krise. Dennoch müssen stets zeit- und kontextbasierte Strategien angewandt werden, um situationsgerecht handeln zu können.

Da es im D-A-CH-Raum grosse Unterschiede sowohl zwischen den Ländern als auch zwischen den Bundesländern hinsichtlich der Aufstellung und Bestimmungen der Schulaufsicht gibt, kann die Rolle der Schulaufsicht nicht allgemeingültig definiert werden. Was sich jedoch als Gemeinsamkeit herauskristallisiert, ist die Entwicklung zu mehr Autonomie der Schulen (Brauckmann et al., 2019; Huber et al., 2016; Brägger et al., 2007; Trachsler, 2004; Wissinger, 2000).

Während der COVID-19 Pandemie variierten zudem die politischen Entscheidungen in Deutschland, Österreich und der Schweiz deutlich. So waren die Schulen in der Schweiz für 6 Wochen geschlossen, während in einigen Bundesländern Deutschlands und Österreichs teilweise über Monate hinweg Fernunterricht durchgeführt wurde. Als Forschungsdesiderata wurden die Erfahrungen und das Management der Auswirkungen der COVID-19 Pandemie durch die Schulaufsicht sowie deren Gemeinsamkeiten und Unterschiede im D-A-CH-Raum identifiziert, wessen sich in der in diesem Paper beschriebenen Studie angenommen wurde.

Zielstellung und Methode: Im Rahmen des Schul-Barometers (Huber et al., 2020) wurden zur Beantwortung der Forschungsfrage "Welche Rolle hat die Schulaufsicht bei der Bewältigung der Auswirkungen der COVID-19-Pandemie auf Schule und Bildung?" semistrukturierte Interviews mit 25 Vertreter*innen der oberen, mittleren und unteren Schulaufsicht aus insgesamt 16 Kantonen / Bundesländern in Deutschland, Österreich und der Schweiz geführt. Die Datenerhebung erfolgte zwischen März und Juni 2021. Durch eine qualitative Inhaltsanalyse nach Mayring (2015) mit deduktiv-induktivem Kodieren konnten die Daten abstrahiert und thematisch strukturiert werden. Ziel war es, in einem explorativen Modus kontextuelle Erfahrungen und aufkommende Probleme zu erruieren und zu verstehen.

Ergebnisse und ihre Bedeutung: Die Studie zeigt auf, welche Herausforderungen, Möglichkeiten, Chancen und Ideen die Schulaufsichten bei der Digitalisierung und Weiterentwicklung des Schulsystems sehen und welche Rolle sie bei der Kompensation von Schereneffekten und beim Umgang mit Risikoschülern spielt. Es zeigten sich drei Phasen (1: Schock und Desorientierung, 2: Verschnaufpause und Zeit zum Planen, 3: Pandemie als kontinuierliche Herausforderung). Als zentrale Aufgaben in der Pandemie kristallisierten sich Kommunikation, erhöhtes Schnittstellenmanagement und mehr Kooperation heras. Zugunsten des Krisenmanagements wurde ausserdem ein zeitweiser Verlust von Pädagogik in Kauf genommen. Im Kontext der Ergebnisse wurden folgende Thesen im Sinne von Empfehlungen von der Schulaufsicht für die Schulaufsicht abgeleitet:

1.) Proaktive und transparente Kommunikation stärken und strategisch ausrichten

2.) Übergeordnete Strategie im Umgang mit Pandemien und anderen Krisen erarbeiten

3.) Beratung und Unterstützung der Schulen ausbauen

4.) Professionalisierung der Schulaufsicht

5.) Gezielt Risikoschüler*innen fördern

6.) Stärkung der Eigenverantwortung von Schule

7.) Kooperation und Verständigung hinsichtlich gemeinsamer Ziele stärken

8.) Spannungsfeld zwischen Eigenverantwortung und zentrale Steuerung meistern

Diese werden bei der GEBF im Sinne des Krisenmanagements in unsicheren Zeiten diskutiert.

A-04: Unterrichtsmerkmale

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S06 S02 B06
Chair der Sitzung: Mario Frei

Merkmale guten Erklärens im Musikunterricht

Mario Frei[1], Sven Hilbert[1], Bernhard Hofmann[2], Gabriele Puffer[2]
[1]Universität Regensburg, Germany; [2]Universität Augsburg, Germany

Theoretischer Hintergrund

Erklären gilt als wichtige Kompetenz von Lehrkräften und wird zu deren Kerntätigkeiten im Unterricht gezählt (z. B. Findeisen, 2017; Fraefel & Scheidig, 2018; Pauli, 2015). Das wirft die Frage nach Qualitätsmerkmalen unterrichtlichen Erklärens auf. Während dazu aus Erziehungswissenschaft und fachdidaktischer Forschung bereits empirische Erkenntnisse vorliegen (z. B. Findeisen, 2017; Kiel, 1999; Kulgemeyer & Tomszyczyn, 2015; Wittwer & Renkl, 2008), zeigt sich die Forschungslage zu Stellenwert, Funktion und Qualität von Erklärhandlungen im Musikunterricht noch defizitär.

Bereits vorgelegt wurde eine erste theoretische Modellierung zu Merkmalen „guten“[1] Erklärens im Fach Musik (Frei et al., 2022). Neben generischen Merkmalen (Strukturiertheit, sprachliche Verständlichkeit, Sprech- und Körperausdruck sowie Adressatenorientierung; Lindl et al., 2019) wurde der Einsatz verschiedener externer Repräsentationsformen (siehe dazu Schnotz & Bannert, 1999) bzw. die sich daraus ergebenden Erklärmodi fachspezifisch konzeptualisiert (optische Repräsentation: z. B. Visualisierungen; akustische Repräsentation: z. B. Hörbeispiele). Im Rahmen des transdisziplinären Forschungsprojekts *FALKE-q* (Schilcher et al., 2021) erfolgte eine empirische Validierung mit Blick auf unterrichtliches Erklären musiktheoretischer Inhalte (Frei, i. Vorb.).

Der Beitrag soll Fragestellung, methodisches Vorgehen und einige Ergebnisse der Teilstudie FALKE-q-Mu vorstellen.

Fragestellung

Die Studie geht von der Annahme aus, dass die Validität potenzieller Merkmale von Unterrichtsqualität perspektivenspezifisch sein kann (vgl. z. B. Kunter & Baumert, 2006; Kleickmann et al., 2019). Für Angehörige verschiedener Statusgruppen (z. B. Schüler:innen, Lehrkräfte) sind möglicherweise unterschiedliche Merkmale einer unterrichtlichen Erklärung qualitätsrelevant. Daraus ergeben sich folgende Fragestellungen:

- Wie unterscheiden sich die Bewertungen von Erklärqualität unter Berücksichtigung von Erklärmodus und Statusgruppe?
- Wie hängen globale Bewertungen von Erklärqualität und Beurteilungen zu den genannten Einzelmerkmalen zusammen? Lassen sich auch hier Unterschiede zwischen den Statusgruppen ausmachen?

Methode

In einem experimentellen Ansatz dienen sieben Video-Vignetten als Stimuli in einem Online-Fragebogen. Jede Vignette enthält eine kurze Erklärung zu insgesamt drei in Curricula der sechsten Jahrgangsstufe typischerweise verankerten musiktheoretischen Inhalten; optische und akustische Formen der Repräsentation werden über alle Videos hinweg systematisch variiert. An der Erhebung nahmen im Fach Musik 271 Personen teil, davon 148 Schüler:innen, 61 Studierende, und 62 Lehrende (41 Lehrkräfte sowie 21 Hochschullehrende). In einem zweistufigen Verfahren bewerteten die Teilnehmenden der drei Statusgruppen jede Erklärung zunächst mit einer Globalnote (als Schulnote von 1 bis 6) und gaben in einem zweiten Durchlauf Beurteilungen zu Einzelmerkmalen

guten Erklärens ab. Die dabei verwendeten sechsstufigen Rating-Skalen basieren auf der theoretischen Konzeptualisierung des Projekts *FALKE-q-Mu*.

Die Mediane der Reliabilitätskoeffizienten (Cronbach's Alpha) pro Skala sind über alle Statusgruppen, Merkmale und Videos hinweg nach Field (2013) und aufgrund der Neukonstruktion als akzeptabel einzustufen (0,62 < a < 0,84).[2] Zur Datenauswertung wurden unter anderem regressions- und varianzanalytische Verfahren verwendet.

Ergebnisse

Eine *mixed ANOVA* zeigte einen großen Effekt des Innersubjektfaktors „Erklärmodus" ($F(2, 518) = 167{,}69$; $p < 0{,}01$; = 0,39). Detailliertere *Post-hoc* Analysen (Werte korrigiert nach *Bonferroni*) ergaben, dass unabhängig der Statusgruppen und unabhängig vom Erklärthema diejenigen Erklärungen signifikant schlechter bewertet werden, in denen optische Repräsentationen fehlen (alle $p < 0{,}01$). Das überrascht vor dem Hintergrund, dass das Fehlen von akustischen Repräsentationen – ein zentrales Spezifikum von Musikunterricht – zu keinen signifikanten Unterschieden in der Erklärqualität führt. Nach Statusgruppen getrennte Regressionsanalysen lassen den Schluss zu, dass für Schüler:innen tatsächlich andere Merkmale relevant sind als für Lehrende und Studierende. Demnach kann die Annahme einer perspektivenspezifischen Validität von Erklärqualität im Fach Musik aufrechterhalten werden.

[1] Der Terminus „gut" wird in Anlehnung an die Kategorie „gutes Unterrichten" nach Berliner (2005) verstanden.

[2] Eine Ausnahme bildet die Skala Sprachliche Verständlichkeit. Die Limitation des niedrigen Median-Werts (Cronbachs-Alpha) von 0,48 wird in der Analyse berücksichtigt.

Wie gut können Schüler:innen historisch denken? Ein Kooperationsprojekt der Geschichtsdidaktik (Universitäten Konstanz und Salzburg), der Empirischen Bildungsforschung (Universität Tübingen) und des Instituts für Qualitätsentwicklung an Schulen Schleswig-Holstein

Christiane Bertram[1], Wolfgang Wagner[2], Christoph Kühberger[3], Nadja Einhaus[4], Karin Hülsen[4], Benjamin Stello[4], Ulrich Trautwein[2]

[1]Universität Konstanz, Germany; [2]Universität Tübingen, Germany; [3]Universität Salzburg, Österreich; [4]Institut für Qualitätsentwicklung an Schulen Schleswig-Holstein (IQSH)

Theoretischer Rahmen

Häufig wird in Gesellschaften mit der Vergangenheit oder der „Geschichte" argumentiert, wenn es um Probleme in der Gegenwart und Lösungsangebote für die Zukunft geht. Um schwierige oder gar unzulässige Umgangsweisen mit Geschichte durchschauen und dekonstruieren zu können, ist in einer demokratisch organisierten Gesellschaft eine historisch-politische Grundbildung unerlässlich. Daher steht neben der Vermittlung relevanten historischen Wissens die Förderung historischer Kompetenzen im Zentrum des Geschichtsunterrichts. Theoretisch fundierte und empirisch erprobte Kompetenzmodelle des historischen Denkens und Lernens (u. a. Gautschi, 2009; Körber et al., 2007) haben in den letzten Jahren Eingang in die Bildungs- und Lehrpläne aller bundesdeutscher Länder gefunden und bestimmen die Aus- und Fortbildung der Lehrkräfte.

Auf dieser Grundlage wurde auch in den Fachanforderungen Geschichte in Schleswig-Holstein (Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein [MSB], 2016) definiert, welche Kompetenzen im Sinne überprüfbarer Standards die Lernenden am Ende der Sekundarstufe I erreicht haben sollen. In einem Kooperationsprojekt des Instituts für Qualitätsentwicklung an Schulen Schleswig-Holstein (IQSH) mit dem Hector-Institut für Empirische Bildungsforschung an der Universität Tübingen und der Fachdidaktik Geschichte an den Universitäten Konstanz und Salzburg haben wir untersucht, inwieweit die Lernenden die für das Ende der Sekundarstufe I formulierten Ziele erreichen.

Forschungsfragen

Wir wollten einen Einblick in den Leistungsstand der Lernenden gewinnen und Unterschiede nach Klassenzugehörigkeit und Schulform überprüfen. Zudem sollten die Informationen über die Testergebnisse der Lernenden den Lehrkräften die Möglichkeit geben, den Leistungsstand ihrer Schüler:innen bezogen auf den Erwerb historischer Kompetenzen zu erfahren. Darüber hinaus wollten wir herausfinden, wie viele Schüler:innen einen Mindeststandard (nicht) erreichen.

Methodisches Vorgehen

Um diese Fragen zu beantworten, führten wir im Herbst 2021 online eine Studie durch, in der erprobte Aufgaben aus einem historischen Kompetenztest, dem sog. HiTCH-Test (Authors, 2017; Authors, 2022) eingesetzt wurden. 111 Klassen und 2.258 Schüler:innen, von denen sich 52% dem weiblichen Geschlecht zugeordnet haben, nahmen an der Erhebung teil. Jeweils 50% waren zum Testzeitpunkt in der achten bzw. neunten Klassenstufe, 26% waren auf einer Gemeinschaftsschule (davon 63% mit Oberstufe), 74% besuchten das Gymnasium. Gemeinsam mit Lehrkräften aus Schleswig-Holstein führten wir einen eintägigen *Standard Setting* Workshop (Tannenbaum & Katz, 2013) durch, in dem wir in einem standardisierten und iterativen Verfahren zu einem gemeinsamen Verständnis der „Minimalstandards“ gelangen wollten: Was sollen Schüler:innen am Ende der Sekundarstufe I mindestens verstanden haben?

Ergebnisse

Zentrale Ergebnisse des Projekts waren:

(1) Der eingesetzte standardisierte Online-Test hat sich aus psychometrischer Sicht als geeignet erwiesen, um historische Kompetenzen reliabel und validie zu erfassen, und fand eine hohe Akzeptanz bei den Schüler:innen und Lehrkräften.

(2) Der Leistungsunterschied zwischen den teilnehmenden Schüler:innen am Gymnasium und der Gemeinschaftsschule war deutlich ausgeprägt und war mehr als doppelt so groß wie der Unterschied zwischen Acht- und Neuntklässler:innen. Außerdem zeigten sich auch nach Kontrolle wichtiger Faktoren wie der Schulform sowie der individuellen Voraussetzungen der Schüler:innen bedeutsame Leistungsunterschiede zwischen einzelnen Schulklassen.

(3) Im Rahmen eines *Standard Setting*-Workshops mit Lehrkräften aus Schleswig-Holstein wurde ein Vorschlag für einen empirisch verankerten Mindeststandard erarbeitet. Ein bzw. zwei Schuljahre vor dem Ende der Sekundarstufe I hatten 21% der Neuntklässler:innen am Gymnasium und 61% an der

Gemeinschaftsschule diesen Mindeststandard (noch) nicht erreicht. In der verbleibenden Schulzeit würde voraussichtlich der Anteil von Schüler:innen, deren Leistungen unterhalb des Mindeststandards liegt, substanziell sinken, doch würden basierend auf einer Projektion immer noch rund 35% der Schüler:innen der Gemeinschaftsschule den Mindeststandard nicht erreichen.

Ausblick

Mit dem Kooperationsprojekt wurde ein Prozess angestoßen, in dem die Forschenden gemeinsam mit den Fachlehrkräften in Schleswig-Holstein an einer Verständigung über Kriterien arbeiten, ab welchen Testergebnissen das historische Denken den in den Fachanforderungen formulierten Erwartungen, den „Standards“, entspricht.

Schriftliches Urteilen im Politikunterricht der Sekundarstufe I - Eine Interventionsstudie zum textsortenbasierten Schreiben im Fachunterricht

Claudia Forkarth
Universität Duisburg-Essen, Germany

Theoretischer Hintergrund

Sprachliches Lernen als Aufgabe aller Fächer zu verstehen und umzusetzen, ist eine zentrale Zielstellung sprachsensibler Schulentwicklung (Roll et al., 2019; Thürmann & Volmer, 2017). Dabei ist die Bedeutung konzeptuell schriftsprachlicher Fähigkeiten für den Schulerfolg und die gesellschaftliche Teilhabe hervorzuheben (OECD, 2006; Reiss et al., 2016).

Am Ansatz der „durchgängigen Sprachbildung“ (Gogolin, 2017) setzt die Arbeit an und zielt auf die Förderung fachspezifischer (schrift-)sprachlicher Fähigkeiten im Politik-/Gesellschaftslehreunterricht am Beispiel der Textsorte *Politisches Urteil*. Die politische Urteilsfähigkeit wird als eine zentrale Dimension der Politikkompetenz (Detjen et al., 2012) verstanden und bietet ebenso wie die Modellerweiterung (Manzel & Weißeno, 2017) einen theoretischen Bezugsrahmen politikdidaktischer Forschung.

Insbesondere mit dem Ziel, Schüler*innen zu mündigen Bürger*innen auszubilden, zeigt sich, wie voraussetzungsreich politische Partizipation und Diskursteilnahme sind und dass Unterstützungsmöglichkeiten für die Lernenden auf unterschiedlichen Ebenen erforderlich sind (Dutz, 2020). Es fehlen schriftsprachliche ebenso wie politikbezogene Kompetenzen, deren Erwerb die Schüler*innen vor große Herausforderungen stellt. Es fällt ihnen einerseits schwer, politische Urteile zu fällen, andererseits wird das schriftliche Urteilen selten gefördert und explizit angebahnt, obwohl Schüler*innen Fachinhalte und Wissen um fachspezifische Sprachstrukturen benötigen, um den Anforderungen des Schreibens im Fachunterricht gerecht zu werden (Oleschko 2012).

Forschungsdesign

Die quasi-experimentelle Interventionsstudie (Pre-Post-Design) untersucht die schriftliche Urteilsfähigkeit von Schüler*innen der Jahrgangsstufen sieben/acht an Gesamtschulen in NRW anhand der sprachlichen Handlung *Begründen* sowie den Effekt einer fachspezifischen, textsortenbasierten Schreibförderung (Genre-Cycle, Taufik, 2009; Callaghan et al. 1993; Roll et al. 2022).

Im Vortrag soll ein Überblick über das abgeschlossene Dissertationsprojekt (Forkarth, im Druck) und ausgewählte Ergebnisse gegeben werden. Die zentrale Forschungshypothese ist, dass eine textsortenbasierte Schreibförderung im Politikunterricht positive Effekte auf das Schreiben politischer Urteile im Fachunterricht hat.

Dabei wird (**1**) die Ausgangslage der Schüler*innen untersucht (*deskriptive Analyse*), (**2**) die Intervention auf ihre Wirksamkeit überprüft (*Varianzanalyse mit Messwiederholung*) und (**3**) Einflussfaktoren auf das fachspezifische Schreiben bestimmt (*Cross-Lagged-Panel-Design, Regressionsanalyse*).

Datenerhebung und -auswertung

Die Daten wurden 2019 an zwei Gesamtschulen in Nordrhein-Westfalen und 13 Klassen mittels Schreibaufgaben und weiteren Testinstrument zur Überprüfung unterschiedlicher Einflussgrößen (u.a. Fachwissenstest, C-Test, SLS) erhoben (9 Interventions-, 4 Kontrollklassen; $N_{Schüler*innen}$ = 358). Die Testverfahren zur Erhebung und Auswertung der Textprodukte der Schüler*innen sind standardisiert, ein Kategoriensystem mit Ankerbeispielen zur Auswertung der Schreibprodukte wurde entwickelt, mehrfach pilotiert und in Doppelrating-Verfahren mit Rater-Schulungen durchgeführt (ICC(A,2) = .779-.988).

Ergebnisse

Ein erster Einblick in die Ergebnisse entlang der Forschungsfragen zeigt:

(1) Ausgangslage: Die Schüler*innen weisen Lücken in der textsortenspezifischen Schreibfähigkeit (*begründen*) auf. Es zeigen sich außerdem Unterschiede in der Textsortenfähigkeit im Deutsch- und Politikunterricht (t1, $t(247) = 6.852$, $p = .000$; t2 $t(225) = 9.523$, $p = .000$). Schüler*innen mit und ohne Zuwanderungshintergrund unterscheiden sich in der Ausgangslage (gemessen an der Textsortenfähigkeit) signifikant voneinander: $t(252) = 2.915$, p = .004. Der Zuwanderungshintergrund korreliert außerdem negativ mit der Textsortenfähigkeit ($r = -.205$**), dem Fachwissen ($r = -.297$**) sowie den allgemeinsprachlichen Fähigkeiten ($r = -.283$**).

(2) Wirksamkeit: Es lässt sich kein Lernzuwachs (unter Berücksichtigung der Kontrollgruppe) der Schüler*innen vom ersten zum zweiten Messzeitpunkt in ihrer Fähigkeit, schriftliche politische Urteile zu verfassen, nachweisen ($F(1, 197) = .633$, $p = .472$) (s.a. Hermansson et al, 2019, zum Nicht-Effekt von Interventionen).

(3) Einflussfaktoren: In der Regressionsanalyse ($|R^2| = .23$) zeigen sich signifikante Effekte des Geschlechts (β = -.178; t(134) = -2.214, p = .029), des Informationsverhaltens (β = .226; t(134) = 2.500, p = .014) und des Fachwissens (β = .232; t(134) = 2.677, p = .008). In der SchriFT II-Studie zeigt sich außerdem, dass weniger die Herkunft als vielmehr fachspezifische und sprachbezogene Faktoren Einfluss nehmen (Krabbe & Enzenbach 2022).

Weitere Ergebnisse z.B. hinsichtlich der Jahrgangsstufen- und Klassenzugehörigkeit, Mehrsprachigkeit oder Fachwissen sowie Limitationen der Studie werden diskutiert.

„Fair Debattieren und Erörtern“ – ein evidenzbasiertes Forschungsprojekt im Feld Schule

Lucas Deutzmann, Winnie-Karen Giera, Sarah Risse

Universität Potsdam, Germany

Ein zentrales Ziel der inklusiven Deutschdidaktik ist es, Schüler:innen der Sekundarstufe I durch die Vermittlung literaler Kompetenzen den Zugang zu Bildungsabschlüssen zu ermöglichen und dadurch Bildungsgerechtigkeit zu erzielen (OECD, 2021). Schulleistungsstudien zeigen, dass Schüler:innen Barrieren im Bildungsverlauf aufweisen, wenn ihnen diese literalen Kompetenzen fehlen (Stanat et al., 2010; Neumann, 2010). Daher gilt es, alle Schüler:innen, und besonders die Schreiber:innen, Sprecher:innen und Leser:innen mit gering ausgeprägten literalen Fähigkeiten, zu fördern, damit diese im beruflichen, kulturellen und politischen Kontext gesellschaftlich partizipieren können. Das mündliche und schriftliche Abwägen von Argumenten eröffnet durch das simultane sprachliche und demokratische Handeln ergiebige Möglichkeiten, literale Kompetenzen und das Bestreben zur politischen und gesellschaftlichen Teilhabe bei Schüler:innen zu fördern (Achour et al. 2020, Winkler, 2005). Trotz dessen lässt sich der Einfluss des mündlichen Argumentierens auf die Fähigkeit zum Schreiben dialektischer Erörterungen im Deutschunterricht am Ende der Sekundarstufe I als Forschungsdesiderat bezeichnen. Dabei ist das Schreiben einer dialektischen Erörterung eine relevante Prüfungsaufgabe im Rahmen des Mittleren Schulabschlusses im Fach Deutsch in der zehnten Jahrgangsstufe.

Auf der Grundlage jenes Forschungsstandes wurde das Unterrichts- und Forschungsprojekt *Fair Debattieren und Erörtern* konzipiert. Dabei handelt es sich um eine kontrollierte, quasi-experimentelle Interventionsstudie im Paneldesign in der Klassenstufe 9 an Gymnasien und Nicht-Gymnasien mit 355 Schüler:innen. In diesem Kontext wird die zentrale Fragestellung untersucht*, wie (schrift-)sprachliche Kompetenzen von Schüler:innen im neunten Jahrgang durch zwei Unterrichtsreihen zu den Themen Debattieren und schriftliches Erörtern erfasst und weiterentwickelt werden können.* Die Studie untersucht hierbei die wechselseitige Beeinflussung der Debattierkompetenz (Treatment D) einerseits und der Schreibkompetenz (Treatment E) andererseits. In diesem Zusammenhang gingen die geschulten Mitglieder des Lehrstuhlteams im Sinne des Design-Based-Research-Ansatzes (Philipakos et al., 2021) in das Feld Schule, um die Intervention in Form der beiden Unterrichtsreihen durchzuführen. Das Projekt wurde in der ersten Phase (11/2021-05/2022) pilotiert und im Sinne des DBR-Ansatzes sowohl innerhalb des Lehrstuhlteams als auch mit Lehrkräften evaluiert, reflektiert und angepasst. Der innovative Charakter der Datenerhebung liegt in der Erfassung der Schreibprozesse als Verlaufsdiagnostik mithilfe der digitalen Erhebungsplattform *Gorilla.sc*. begründet, welche zu vier Messzeitpunkten (Pre-, Zwischen-, Post- und Follow-up-Testung) in randomisierten Interventionsgruppen sowie in Kontrollgruppen erfolgte. Die Messung der Schreibkompetenz wurde in diesem Zusammenhang durch vier Schreibaufgaben (Textsorte dialektische Erörterung) ausgewählter MSA-Prüfungen der letzten zehn Jahre realisiert. Alle anonymisierten Erörterungstexte (n =1000) wurden, angelehnt an das IMOSS-Kodierverfahren (Neumann & Matthiesen, 2011), im Double-Blind-Verfahren holistisch und analytisch codiert (Cohen’s Kappa =.727). Die Datenaufbereitung erfolgte durch R Studio in Form von Mehr-Ebenen-Analysen.

Der Vortrag fokussiert die Ergebnisse der Interventionsstudie hinsichtlich der Schreibkompetenz der Schüler:innen in Form der Textlänge, der Textqualität sowie der digital erfassten Schreibprozesse. Die Ergebnisse zeigen hinsichtlich der Reihenfolge der Treatments, dass die Texte derjenigen

Schüler:innen, welche zuerst debattierten und danach schriftlich erörterten, eine größere Textlänge- und Qualität aufwiesen als in der anderen Interventionsgruppe sowie der Kontrollgruppe. In diesem Zusammenhang konnten die Textlänge, die Anzahl der Wörter im (digitalen) Schreibplan, die Zeit im (digitalen) Schreibfeld, die Zeit im Schreibplan und die Anzahl der Wechsel zwischen Schreibfeld und Schreibplan als hoch signifikante Prädiktoren für die Textqualität ermittelt werden ($p<0.01$). Zudem bestand ein höchst signifikanter Zusammenhang zwischen den vier Messzeitpunkten und allen abhängigen Variablen in allen Probandengruppen ($p<0.001$). Die Ergebnisse verdeutlichen die Relevanz von Interventionsstudien im Längsschnitt im DBR-Design, um Transferprozesse zwischen Universität und Unterrichtspraxis zu initiieren.

A-05: Sprache im Unterricht

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S06 S03 B06
Chair der Sitzung: Robert Gieske

Vergleichende Analyse der Wirksamkeit zweier sprachsensibler Unterrichtsansätze beim Chemielernen

Robert Gieske, Sabine Streller, Claus Bolte
Freie Universität Berlin, Deutschland

Theoretischer Hintergrund

Sprachliche Kompetenzen von Schüler*innen üben nachweislich einen Einfluss auf den Lernerfolg in den naturwissenschaftlichen Fächern aus (Bird & Welford, 1995; Gogolin & Lange, 2011). Aus diesem Grund wurden bereits zahlreiche Ansätze, wie das Scaffolding (Hammond & Gibbons, 2005), das Translanguaging (García, 2011) oder das Sheltered Instruction Observation Protocol (SIOP, Echevarria et al., 2006), entwickelt und größtenteils erfahrungsbasiert erprobt. Es liegen bislang kaum empirische Befunde z.B. aus systematisch angelegten Interventionsstudien vor und die wenigen Arbeiten haben selten eindeutige Belege für die Wirksamkeit der Ansätze hervorgebracht (Woerfel et al., 2020, 8f.). Wir wollen mithilfe unserer Arbeit zur Förderung fachbezogener Sprachkompetenzen von Schüler*innen einen Diskussionsbeitrag leisten, indem wir Ergebnisse einer empirisch angelegten Vergleichsstudie unter Anwendung des Disaggregate-Instruction-Ansatzes (kurz DIA) von Brown et al. (2010) vorstellen.

In Studien mit mehrsprachigen Schüler*innen im Fach Science konnten Brown et al. (2010) nachweisen, dass Lernende, die mithilfe des DIA unterrichtet wurden, ein besseres konzeptuelles Verständnis als Schüler*innen der Kontrollgruppe erwarben und dieses anschließend im Vergleich überlegen kommunizieren konnten (Brown et al., 2010). Die Besonderheit des DIA besteht darin, dass neu zu erlernende Fachtermini im Unterricht zunächst zurückgehalten und die fachlichen Konzepte in einer für die Schüler*innen bekannten Sprache vermittelt werden. Erst wenn diese ein grundlegendes Verständnis zum Thema herstellen konnten, werden die korrespondierenden Fachtermini eingeführt und gefestigt (Brown et al., 2010, 1474f.). Diese Trennung von konzeptuellem und fachsprachlichem Lernen soll Überforderung auf Seiten der Lernenden vorbeugen und unterstützt Schüler*innen, die begrenzten Kapazitäten des Arbeitsgedächtnisses zielgenau einzusetzen (Brown et al., 2019).

Fragestellung

Um die Wirksamkeit des Ansatzes zu untersuchen, gehen wir folgender Forschungsfrage nach: In welchem Maße profitieren Lernende mit divergenten sprachlichen Kompetenzen beim Erwerb chemischen Fachwissens, wenn sie im Sinne eines sprachsensiblen Disaggregate-Instruction-Ansatzes im Fach Chemie unterrichtet werden?

Methode

Zu diesem Zweck haben wir eine Interventionsstudie im Prä-Post-Design für den Chemieunterricht der Jahrgangsstufen 8/9 an Gymnasien und Integrierten Sekundarschulen (ISS) konzipiert. Die Art des Unterrichtsansatzes fungiert als unabhängige Variable und der fachbezogene Lernzuwachs der Schüler*innen als abhängige Variable. Als Intervention dient die Unterrichtsreihe „Das Tote Meer stirbt“, welche sprachsensibel unter Anwendung von Scaffolding (Hammond & Gibbons, 2005) für Interventions- und Kontrollgruppe entwickelt wurde. Zusätzlich zum Scaffolding wurde der Unterricht

für die Interventionsgruppe an den Gestaltungskriterien des DIA orientiert. Um die fachwissenschaftlichen Kenntnisstände der Schüler*innen und deren Lernzuwächse nachzeichnen zu können, nutzen wir einen selbst entwickelten Fachwissenstest mit 16 Items im MC-Format. Außerdem erfassen wir den Sprachstand der Lernenden mithilfe eines erprobten C-Tests (ifbq Hamburg, 2008). Im Zuge unserer Analysen teilen wir die Gesamtstichprobe in Teilstichproben; nämlich einerseits anhand des C-Test-Medians in sprachlich schwächere und sprachlich stärkere Schüler*innen und andererseits in Schüler*innen des Gymnasiums sowie der ISS; jeweils in der Interventions- als auch in der Kontrollgruppe.

Ergebnisse und ihre Bedeutung

Insgesamt konnten Datensätze von 228 Schüler*innen verwendet werden.

Die Ergebnisse im Fachwissenstest verdeutlichen, dass die Schüler*innen einen signifikanten Lernzuwachs von ΔM=5,3 erreichen (p<,001). Dieser fällt für die Lernenden der Interventionsgruppe (n=113, ΔM=5,4) und der Kontrollgruppe (n=115, ΔM=5,1) ähnlich aus. Betrachtet man lediglich Schüler*innen an ISS (n=117) und unterteilt diese anhand des Medians der erreichten Punkte im C-Test (Md=66), so schneiden die sprachlich schwächeren Schüler*innen in der Interventionsgruppe statistisch signifikant besser ab (ΔM=5,4) als die sprachlich schwächeren Lernenden in der Kontrollgruppe (ΔM=3,6). Sprachlich schwache Lernende der Interventionsgruppe erreichten außerdem vergleichbare Lernzuwächse wie sprachlich starke Schüler*innen unter beiden Bedingungen.

Die Ergebnisse belegen, dass durch eine bewusste Trennung des konzeptuellen und fachsprachlichen Lernens eine vulnerable Gruppe in besonderem Maße gefördert werden kann. Bemerkenswert ist, dass sprachlich starke Lernende vergleichbar vom DIA profitieren und darüber hinaus auf sehr ähnliche Weise einen Nutzen ziehen wie sprachlich starke Lernende im sprachsensiblen Unterricht ohne strikte Trennung zwischen konzeptuellem und fachsprachlichem Lernen.

Sprachmittlung im Englischunterricht – eine empirische Untersuchung

Dorothee Kohl-Dietrich[1], Axel Zinkernagel[2]
[1]PH Karlsruhe; [2]Universität Koblenz-Landau

Theoretischer Hintergrund

Sprachmittlung (*mediation*) wurde erstmals im Gemeinsamen Europäischen Referenzrahmen für Fremdsprachen (CEFR 2001) als *language activity* neben *reception*, *production* und *interaction* erwähnt. In der Folge wurde in Deutschland dieser Aufgabentypus in die Bildungsstandards (2003, S. 14; 2012, S. 18) explizit als „fünfte Fertigkeit“ (*fifth skill*) respektive als funktionale kommunikative Kompetenz aufgenommen und verstärkt Sprachmittlungsaufgaben in Lehrwerken und Unterrichtsmaterialien implementiert. Im CEFR *Companion Volume with New Descriptors* (2017) wurde das ursprüngliche Konstrukt Sprachmittlung schließlich umfassend erweitert (vgl. *linguistic, social, pedagogical, cultural mediation*) und um Deskriptoren ergänzt. Eine dieser Subfacetten, cultural mediation, steht im Fokus dieser empirischen Untersuchung. Bei Sprachmittlungsaufgaben liegt im Unterschied zur klassischen Übersetzung der didaktische Fokus auf einer sinn-, situations-

und adressatengemäßen Übertragung eines Inhalts von einer sprachlichen Varietät in eine andere unter Berücksichtigung des (inter)-kulturellen Kontexts. Dies setzt bei Sprachmittler*innen die Fähigkeit voraus, zwischen unterschiedlichen sprachlichen Varietäten, Registern und kulturellen Kontexten zu agieren. So kann angenommen werden, dass erfolgreiche Sprachmittlung eines hohen Maßes an Sprach- und Kulturbewusstheit bedarf (Byram 2015), weswegen in der Fremdsprachendidaktik vielfach das Potential von Sprachmittlungsaufgaben im Hinblick auf die Förderung trans-/ interkultureller Kompetenzen hervorgehoben wird (z.B: Caspari & Schinschke, 2010; Hallet, 2008; Kolb, 2016; Rössler & Reimann, 2013). Trotz dieser sprachenpolitischen und didaktischen Relevanz der Sprachmittlung gibt es bis dato nur theoretisch deduzierte Modelle von Sprachmittlungskompetenz (Hallet, 2008; Kolb, 2016). Empirische Untersuchungen stellen noch weitestgehend ein Desiderat dar.

Design der Studie und Auswertungsmethode

In diesem Paper werden Erkenntnisse aus einem empirischen Forschungsprojekt vorgestellt, das in den Klassenstufen 10-13 des Gymnasiums (*N* = 92) durchgeführt wurde. So wurden in 3 x 45 Minuten drei Testlets mit drei unterschiedlichen Aufgabenformaten zur Sprachmittlung eingesetzt, um Sprachmittlungskompetenz performativ als abhängige Variable (AV) zu testen. Auf der Grundlage von Regressionsmodellen untersuchte die Studie,

- ob Sprachbewusstheit, Sprachkompetenz und interkulturelle Kompetenz (IKK) Prädiktoren für die Performanz in den Sprachmittlungsaufgaben sind
- welche weiteren interindividuellen Variablen (z.B. Geschlecht, Persönlichkeitseigenschaften, Auslandserfahrung, Mehrsprachigkeit) im Modell zu berücksichtigen sind.

Zur Messung der Konstrukte Sprachbewusstheit, Sprachkompetenz, interkulturelle Kompetenz (IKK) und soziodemographischer Daten als UVs wurden für die DESI-Studie entwickelte Instrumente (z.B. DESI-IKK allgemein und spezifisch auf Großbritannien bezogen, Cloze Tests zur Erhebung des Sprachstandes, Test zur Messung von Sprachbewusstheit-Soziopragmatik, vgl. Wagner et al. 2003) sowie das BIG Five Inventory (BFI-10) zur Messung von Persönlichkeitsvariablen verwendet (Rammstedt et al. 2013). In der DESI-Studie konnte ein deutlicher positiver Zusammenhang zwischen IKK und Gesamtleistung im Englischtest sowie im Soziopragmatik-Test (Nold & Rossa 2007) gefunden werden, auch nach Kontrolle individueller Merkmale (Hesse & Göbel 2007, Hesse, Göbel, & Jude 2008). Zur Messung der Sprachmittlungskompetenz als AV wurden Testaufgaben auf der Basis theoretischer Modelle zur Sprachmittlung entwickelt, die Sprachproduktion erfordern und zu deren Lösung pragmatisches und kulturelles Wissen notwendig ist oder in denen kulturspezifische Konzepte gemittelt werden sollen (vgl. auch Kohl-Dietrich 2019). Dadurch soll Konstruktvalidität gewährleistet werden. So wurden insgesamt drei verschiedene Aufgabenformate entwickelt: MT1: Mediating a culturally specific concept („Brezelaufgabe“), MT2: Mediating a political text (excerpt from D. Hayali, *Haymatland* (2018)), MT3: Intercomprehension Activity/ Mediating a tourist flyer from Dutch into German.

Die Daten wurden u.a. mittels multipler Regressionen und Korrelationsanalysen ausgewertet. So haben sich z.B. in Bezug auf Aufgabentyp MT1 Sprachbewusstheit und Sprachkompetenz als signifikante Prädiktoren gezeigt. In Bezug auf Aufgabentyp MT2 waren Sprachkompetenz, Extraversion und Neurotizismus signifikante Prädiktoren. Korrelationsanalysen spiegeln eine mittlere

signifikante Korrelation der Aufgabentypen MT1 bis 3 wider. Dies kann als Hinweis auf Konstruktvalidität interpretiert werden in dem Sinne, dass alle drei Aufgaben unterschiedliche Facetten von Sprachmittlung messen. Neben der Darstellung weiterer Ergebnisse der Untersuchung diskutiert der Beitrag kritisch Herausforderungen der empirischen Messung komplexer Konstrukte wie interkulturelle Kompetenz.

Beziehungen zwischen Sprachreflexionsfähigkeiten und anderen sprachlichen sowie allgemein-kognitiven Fähigkeiten

Thomas Canz, Kathrin Jonkmann
FernUniversität in Hagen, Deutschland

Die Fähigkeit, die Systematik und Struktur von Sprache sowie ihre Verwendungsweisen zu reflektieren ist eines der zentralen Ziele im Deutschunterricht, die in den Bildungsstandards unter dem Kompetenzbereich „Sprache und Sprachgebrauch untersuchen“ verankert ist (Feilke & Jost, 2015; KMK, 2004; Oomen-Welke & Bremerich-Vos, 2014). Dieser Kompetenzbereich zieht sich zum einen durch alle anderen für das Fach Deutsch festgelegten Kompetenzbereiche, steht aber gleichermaßen als Metakompetenz über ihnen, da hier metasprachliche Fähigkeiten wie bspw. das Wissen über Sprache die zentrale Rolle einnehmen. Sprachreflexionsfähigkeiten können als Teil des Konstrukts *Language Awareness* aufgefasst werden. Es umfasst explizites Wissen über Sprache und die bewusste Wahrnehmung und Sensibilität beim Sprachenlernen, Sprachunterricht und Sprachgebrauch, bspw. die Rolle von Sprache beim Denken, beim Lernen, im sozialen Leben oder die Beziehung zwischen Sprache und Kultur (Association of Language Awareness, 2012; van Lier, 1995).

Im Vergleich zu anderen Kompetenzbereichen ist dieser Kompetenzbereich bisher in der empirischen Bildungsforschung kaum erforscht; Untersuchungen bei Erwachsenen fehlen gänzlich. Ziel dieser Studie ist es, diese Forschungslücke zu schließen und der Frage nach dem nomologischen Netzwerk von Sprachreflexionsfähigkeiten und den Relationen des Konstrukts zu anderen sprachlichen Fähigkeitskonstrukten, i.e. Lesekompetenz und Wortschatz, sowie zum logischen Denken nachgegangen.

Im Rahmen einer Online-Studie nahmen 481 Personen (69.5 % weiblich, 30.5 % männlich, 0.5 % divers) im Alter zwischen 17 und 77 Jahren (M = 37.4; SD = 12.6) teil. Die Teilnehmenden bearbeiteten in einer 90-minütigen Untersuchung diverse Fragebögen sowie Tests zur Sprachreflexionsfähigkeit, zum Lesen, zum Wortschatz und zum logischen Denken. Zur Erfassung von Sprachreflexionsfähigkeiten wurden Aufgaben, die am Institut zur Qualitätsentwicklung im Bildungswesen (IQB) entwickelt und unter psycholinguistischer Expertise ergänzt und für den Einsatz bei Erwachsenen adaptiert wurden, eingesetzt (55 Items). Als Lesetest fungierte (approximativ) eine Leseverständnisaufgabe des IQB (18 Items). Der Wortschatz wurde mit dem WST (Schmidt & Metzler, 1992) erfasst (42 Items). Das logische Denken wurde in seinen Subdomänen *verbal* durch Subtests zu sprachlichen Analogien (16 Items) und Syllogismen (12 Items) sowie *figural* durch Subtests zu Mustergruppen (10 Items), Figurenfolgen (8 Items) und Figurenmatrizen (10 Items) erfasst. Für alle Tests war die Bearbeitungszeit limitiert.

IRT-Modellierungen der Konstrukte zeigten zufriedenstellende bis sehr gute Model-Fit-Werte und EAP-Reliabilitäten (> .80). Konstruktkorrelationen lagen für Sprachreflexionsfähigkeiten mit Lesen bei *r* = .53, mit Wortschatz bei *r* = .55 und mit logischem Denken bei *r* = .60 (bzw. bei *r* = .63 mit verbalem und bei *r* = .51 mit figuralem logischem Denken).

Explorative Faktorenanalysen zeigten, dass sich die Konstrukte in einem zweidimensionalen Raum verorten lassen, wobei sich die Dimensionen in etwa als kristalline und fluide Fähigkeitsdimensionen interpretieren lassen. Der Wortschatz wies hierbei mehrheitlich kristalline Anteile, das logische Denken (vor allem die figurale Dimension) dominant fluide Anteilen auf. Sprachreflexionsfähigkeiten und Lesen zeigten mittlere Anteilen an beiden Dimensionen.

Die Ergebnisse stehen im Einklang mit bisherigen Befunden, dass sprachliche Konstrukte untereinander hoch korrelieren (u. a. Canz, 2015; Jude, 2008), verdeutlichen aber auch ihre empirische Trennbarkeit untereinander sowie deren Trennbarkeit von allgemein-kognitiven Fähigkeiten (u. a. Schipolowski et al., 2016).

Entwicklung und Erprobung einer Lernverlaufsdiagnostik Sprache

Melanie Besca, Marco Ennemoser
Pädagogische Hochschule Ludwigsburg, Germany

Theoretischer Hintergrund. Verfahren zur Diagnostik des Sprachstands sind meist zeitintensiv, umfangreich und nicht für einen engmaschig wiederholten Einsatz konzipiert. Verfahren der Lernverlaufsdiagnostik[1,2] (LVD) dauern meist wenige Minuten, enthalten Parallelformen, sind änderungssensitiv und konzentrieren sich meist auf einen geeigneten „Globalindikator“ als Kompetenzmaß. Bisher gibt es in der LVD keine Verfahren im Bereich der produktiven sprachlichen Fähigkeiten.

Zu den wichtigsten Vorteilen der LVD zählt, dass man sog. Nonresponder früher als üblich entdecken kann. Nonresponder sind Kinder, die nicht oder zu wenig von einer Fördermaßnahme profitieren. Sie fallen meist zu spät auf, z.B. wenn sie in der Schule scheitern. Langfristig sind sie dadurch in ihrer Teilhabe beeinträchtigt. Durch den Einsatz von LVD könnte der individuelle sprachliche Lernfortschritt kontinuierlich evaluiert werden. So kann rechtzeitig erkannt werden, ob die Entwicklung stagniert, also ein Kind nicht von einer Förderung profitiert und die Fördermaßnahme modifiziert oder intensiviert werden muss.

Das Nachsprechen von Sätzen bzw. Satzgedächtnis ist ein aussagekräftiges Maß in der Sprachdiagnostik z.B. im SETK 3-5[3] und dient der Identifikation einer Sprachentwicklungssörung[4] bei ein- und mehrsprachigen Kindern[5] oder als Prädiktor für Lese- und Rechtschreibfähigkeiten[6]. Es stellt Anforderungen an das phonologische Arbeitsgedächtnis, das semantische und syntaktische Wissen[7]. Zudem differenziert es über einen größeren Altersbereich hinweg. Somit scheint das Nachsprechen von Sätzen ein vielversprechendes Globalmaß sprachlicher Fähigkeiten zu sein und sich auch ökonomisch für den Einsatz im Rahmen der LVD zu eignen.

Fragestellung. Ziel ist die Entwicklung und Erprobung eines Verfahrens zur Darstellung des Entwicklungsverlaufes produktiver sprachlicher Fähigkeiten und die Klärung, ob sich das Nachsprechen von Sätzen hierfür eignet.

Methode. Insgesamt wurden elf Paralleltests des Sätze Nachsprechens mit jeweils 16 Sätzen entwickelt (vier sinnvolle und zwölf sinnlose, aber grammatisch korrekte Sätze). Die Wörter wurden nach Zufallsprinzip aus der childLex-Datenbank[8] ausgewählt. Die Konstruktion der Sätze orientiert sich am Erwerb der unauffälligen Grammatikentwicklung[9]. Die Sätze haben einen ansteigenden Schwierigkeitsgrad mit zunehmender Satzlänge und grammatikalischer Komplexität.

Die Paralleltests wurden 2021/ 2022 im Quer- und Längsschnitt mit jeweils ca. 100 Kindern im Kindergarten sowie am Sonderpädagogischen Bildungs- und Beratungszentrums mit Förderschwerpunkt Sprache erprobt. Die Daten zum Längsschnitt werden aktuell ausgewertet.

Zur Erprobung der Vergleichbarkeit und Reliabilität wurden alle Paralleltests mit jedem Kind innerhalb von zwei Wochen (MW 13 Tage) durchgeführt. Mit der Erprobung im Längsschnitt soll die Frage nach Validität und Änderungssensitivität des Verfahrens beantwortet werden. Über einen Zeitraum von sechs Monaten wurde alle vier Wochen ein Paralleltest durchgeführt. Als Kontrollvariablen wurden Prä/ Post konstruktnahe (z.B. Wortschatz, Grammatik, Arbeitsgedächtnis) sowie konstruktferne Tests durchgeführt (z.B. Begabung und mathematische Fähigkeiten).

Ergebnisse und ihre Bedeutung.

Lernverlaufsdiagnostik, Sprachdiagnostik, Sätze Nachsprechen

Die Erprobung der Vergleichbarkeit der Paralleltests anhand einer Varianzanalyse mit Faktor Paralleltest ergab keinen signifikanten Unterschied der Mittelwerte $F(21,828)=,366$; $p=.996$[10]. Daraus kann abgeleitet werden, dass die Paralleltests vergleichbar schwer und somit die Voraussetzung für die Erprobung im Längsschnitt erfüllt sind. Die Berechnung der Reliabilität mittels Cronbachs Alpha ergaben sehr gute interne Konsistenzen zwischen .92. und .94 und sehr gute Trennschärfen von .46 bis .86[10]. Die Berechnungen zur konvergenten Validität zeigen insgesamt eine hohe Übereinstimmung der neu entwickelten Paralleltests des Sätze Nachsprechens mit Testergebnissen, die das gleiche Konstrukt messen oder theoriebegründet zusammenhängen, z.B. Sätze Nachsprechen aus SETK 3-5 ($n=51$) $r= .82 - .93$, $p < .01$[10].

Vorläufige Auswertungen zur Änderungssensitivität zeigen signifikante Anstiege der Mittelwerte im Verlauf von sechs Monaten und deuten darauf hin, dass Änderungen in der sprachlichen Entwicklung mit den Paralleltests messbar sind. Ende des Jahres werden diese Daten vollständig vorliegen.

Die bisherigen Erprobungsergebnisse deuten darauf hin, dass sich die neu entwickelten Paralleltests zum Sätzen Nachsprechen eignen, um die sprachliche Entwicklung i.S. einer Verlaufskontrolle systematisch und ökonomisch erfassen zu können. Dadurch können Nonresponder frühzeitig erkannt und Fördermaßnahmen angepasst werden, um langfristig ggf. Lernstörungen präventiv entgegenzuwirken.

A-06: Berufswahl

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S06 S04 B06
Chair der Sitzung: Jan Hofmann

Nach den Berufswahlpraktika ist vor der Wahl des Ausbildungsberufs: Vorhersage des Entscheids für die Weiterverfolgung eines Praktikumberufs als potenzieller Ausbildungsberuf im Berufswahlprozess

Jan Hofmann, Markus P. Neuenschwander
Fachhochschule Nordwestschweiz, Schweiz

Berufswahlpraktika haben eine wichtige Rolle im Berufswahlprozess. Sie helfen, Unsicherheiten zu reduzieren und das eigene Selbstbewusstsein zu stärken (Beinke, 2013). Jugendliche erhalten ausserdem die Möglichkeit, ihre Wunschberufe und angestrebten Berufsfelder auszuprobieren (Neuenschwander et al., 2018). Häufig absolvieren Jugendliche mehrere Praktika, wobei unklar ist, unter welchen Bedingungen ein Praktikumberuf im Berufswahlprozess weiterverfolgt und unter welchen Umständen vom Beruf abgekehrt wird. Um diese Forschungslücke anzugehen, wurde folgende Fragestellung untersucht: Unter welchen Bedingungen entscheiden sich Jugendliche für die Weiterverfolgung eines Praktikumberufs als potenzieller Ausbildungsberuf im Berufswahlprozess?

Als theoretische Grundlage wurde Gottfredsons (2002) Theory of Circumscription, Compromise, and Self-Creation (TCCSC) verwendet. Für das Jugendalter besagt die Theorie, dass die Zugänglichkeit eines Berufs (z. B. Stellenangebot) berufswahlbezogene Entscheide am stärksten beeinflusst. Anforderungen, die ein Beruf an eine Person stellt, haben in einem Kompatibilitätsbezug zu den beruflichen Interessen (Übereinstimmung zwischen Anforderungen und Interessen) ebenfalls einen Einfluss auf den Entscheid.

Gestützt auf diese theoretische Grundlage wurde die Forschungsfrage mit Daten einer Schweizer Längsschnittstudie überprüft. Verwendet wurde ein Sample aus 128 weiblichen und 202 männlichen Jugendlichen, die im siebten (T1) und neunten Schuljahr (T2) Fragebögen ausfüllten. Die von den Jugendlichen ausprobierten Praktikaberufe wurden zu T2 erhoben und in einem Mehrebenenmodell auf Level 1 als Gelegenheiten spezifiziert; die Jugendlichen wurden auf Level 2 ins Modell einbezogen. Als abhängige Variable wurde der Entscheid für die Weiterverfolgung eines Praktikumberufs spezifiziert. Die Variable wurde mit der folgenden Frage zu T2 erhoben: «Hat eine der Praktika dazu geführt, dass du dich für/gegen diesen Beruf entschieden hast?» (Antwortoptionen: 1 [*dagegen*], 2 [*weder noch*], 3 [*dafür*]). Die Jugendlichen gaben ihre Weiterverfolgungsentscheide pro Praktikumberuf an. Das Lehrstellenangebot zu den Praktikaberufen (T2) als Zugänglichkeitsaspekt und die Person-Beruf-Kompatibilität zwischen den beruflichen Interessen der Jugendlichen und den beruflichen Anforderungen der ausprobierten Praktikaberufe wurden als unabhängige Variablen ins Modell einbezogen. Für die Operationalisierung des Lehrstellenangebots wurden Daten des Nahtstellenbarometers (Golder & Mousson, 2019, 2020, 2021) verwendet. Die Person-Beruf-Kompatibilität wurde über Angaben der Jugendlichen zu ihren beruflichen Interessen (T1) und Daten zu beruflichen Anforderungen von Berufen (Schmidt, 2008) gebildet. Sowohl die beruflichen Interessen als auch die beruflichen Anforderungen wurden in Anlehnung an Hollands (1997) RIASEC-Berufsklassifizierungssystem entlang der sechs Interessens-/Umweltbereiche handwerklich-technisch, untersuchend-forschend, künstlerisch-kreativ, erziehend-pflegend, führend-verkaufend und ordnend-verwaltend operationalisiert. Pro Interessens-/Umweltbereich wurde eine Person-Beruf-

Kompatibilitätsvariable erstellt, indem der z-standardisierte Interessenswert vom z-standardisierten Anforderungswert subtrahiert wurde. Mit Blick auf den Forschungsstand (Taylor & Pryor, 1985) wurde der Einfluss der Person-Beruf-Kompatibilitätsvariablen auf den Weiterverfolgungsentscheid geschlechtermoderiert geprüft.

Das Mehrebenenmodell zeigte, dass sich Jugendliche umso wahrscheinlicher für die Weiterverfolgung eines Praktikumberufs im Berufswahlprozess entscheiden, je grösser das Lehrstellenangebot für diesen Beruf ist. Weiter stellte sich ausschliesslich bei den weiblichen Jugendlichen heraus, dass das Ausmass führend-verkaufender Anforderungen eines ausprobierten Praktikumsberufs den Weiterverfolgungsentscheid ebenfalls signifikant beeinflusst, wobei das Ausmass des Kompatibilitätsbezugs zu den eigenen führend-verkaufenden Interessen irrelevant ist. Je höher die führend-verkaufenden Anforderungen eines ausprobierten Praktikumberufs sind, desto grösser ist die Wahrscheinlichkeit, dass sich weibliche Jugendliche gegen die Weiterverfolgung dieses Berufs entscheiden. Der Einfluss der restlichen Person-Beruf-Kompatibilitätsvariablen auf den Weiterverfolgungsentscheid war insignifikant (sowohl in einem Kompatibilitätsverständnis als auch als berufliche Anforderungen ohne Kompatibilitätsbezug).

Die Ergebnisse stützen die in der TCCSC postulierte Zugänglichkeitsthese, wonach Personen Berufe mit grösserer Wahrscheinlichkeit weiterverfolgen, wenn für diese eine gute Zugänglichkeit besteht (z. B. im Sinne eines grossen Lehrstellenangebots). Eine mögliche Erklärung für den negativen Einfluss führend-verkaufender Anforderungen auf den Weitverfolgungsentscheid ist vor dem Hintergrund des in der TCCSC enthaltenen Kompatibilitätsgedankens, dass weibliche Jugendliche ihre führend-verkaufenden Fähigkeiten als zu gering einschätzen und Berufe mit hohen führend-verkaufenden Anforderungen als für sich inkompatibel wahrnehmen. Sollten sich die Ergebnisse in zukünftigen Studien replizieren lassen, könnte die Förderung des führend-verkaufenden Fähigkeitsselbstkonzepts bei weiblichen Jugendlichen eine mögliche Massnahme zur Erreichung eines positiven Weiterverfolgungsentscheids und damit einer potenziellen Fortsetzung des Wunschberufs darstellen.

Direkter Einstieg in die berufliche Grundbildung – eine Frage von Selbstwert oder guten Noten?

Sara Hutchison, Lars Balzer
Eidgenössische Hochschule für Berufsbildung EHB, Switzerland

In der Schweiz müssen junge Menschen schon in relativ jungen Jahren erste Entscheidungen über ihre berufliche Zukunft treffen. Die obligatorische Schulzeit in der Schweiz umfasst neun Jahre (ohne Kindergarten). Fast die Hälfte (48%) aller Schülerinnen und Schüler treten am Ende der obligatorischen Schulzeit in eine berufliche Grundbildung ein (gfs.bern, 2021). 40% verfolgen einen allgemeinbildenden Weg und müssen sich deshalb noch nicht für einen Beruf entscheiden. Die restlichen 12% verfügen über keine direkte Anschlusslösung nach der obligatorischen Schulzeit. Von diesen 12% entscheiden sich 9% für ein sogenanntes "Brückenangebot", die anderen absolvieren entweder Praktika oder Sprachaufenthalte. Nur eine kleine Minderheit hat keine Pläne (gfs.bern, 2021).

Studien haben gezeigt, dass es unterschiedliche Gründe dafür gibt, weshalb Jugendliche nicht direkt nach der obligatorischen Schulzeit eine Lehre beginnen. Es kann sein, dass sie noch keine Lehrstelle

gefunden haben, dass sie sich noch nicht für einen Beruf entschieden haben oder dass sie sich noch nicht bereit für einen "richtigen" Job fühlen (Brahm et al., 2012; Künzli & Scherrer, 2013). Während einige Artikel die Bedeutung von Noten für einen erfolgreichen Direkteinstieg in die berufliche Grundbildung betonen (Buchholz et al., 2012; Fossati et al., 2020), haben andere Untersuchungen Hinweise darauf gefunden, dass es weniger die schulischen Leistungen oder fehlende Kompetenzen sind, die den Unterschied ausmachen (Bayard Walpen, 2013), sondern eher vage berufliche Ziele (Sacchi & Meyer, 2016). Vage berufliche Ziele können ein Indikator für berufliche Unentschlossenheit sein. Studien zu beruflicher Unentschlossenheit und beruflichen Entscheidungsschwierigkeiten fanden einen starken negativen Zusammenhang mit dem Selbstwert (Marcionetti, 2014; Rossier et al., 2022; Udayar et al., 2020; Vasalampi et al., 2010). Es hat sich wiederholt gezeigt, dass zwei weitere Variablen mit dem Selbstwert zusammenhängen, nämlich Persistenz und Autonomie. Menschen mit hohem Selbstwert neigen dazu, in schwierigen Situationen mehr Persistenz zu zeigen (Baumeister et al., 2003; Crocker et al., 2006). Auch Autonomie hängt nachweislich mit dem Selbstwertgefühl zusammen (Hodgins et al., 2007; Kiang & Bhattacharjee, 2019).

Unsere Studie zielt daher darauf ab, folgende Forschungsfrage zu beantworten: Inwiefern ist der Selbstwert, direkt wie auch indirekt via berufliche Zielklarheit, prädiktiv für einen direkten Einstieg in die berufliche Grundbildung? Datengrundlage ist eine Stichprobe von 354 Schweizer Schülerinnen und Schülern, die im Rahmen eines Längsschnittprojektes im 7., 8. und 9. Schuljahr Online-Fragebögen ausgefüllt haben. Das Strukturgleichungsmodell wurde in R mit lavaan gerechnet.

Das resultierende Modell (siehe pdf) zeigt einen guten Fit (robuster RMSEA = .05, robuster CFI = .95, TLI = .94). Der Selbstwert in der 7. Klasse steht sowohl mit Autonomie als auch mit Persistenz in Zusammenhang. Er ist auch stark prädiktiv für Selbstwert und Zielklarheit ein Jahr später, in der 8. Klasse. Diese beiden Variablen sind wiederum prädiktiv für den direkten Einstieg in die berufliche Grundbildung am Ende der 9. Klasse, wobei sich Selbstwert als stärkerer Prädiktor erweist als Zielklarheit. Die Mathematik- und Deutschnoten der Schülerinnen und Schüler in der 8. Klasse (dies ist der Zeitpunkt, an dem die Schülerinnen und Schüler beginnen, sich für eine Lehrstelle zu bewerben) haben hingegen keinen Prädiktionswert für den direkten Einstieg in die berufliche Grundbildung nach der obligatorischen Schulzeit. Unsere Ergebnisse bestätigen Forschungsbefunde, die die Bedeutung nicht-fachlicher Eigenschaften für einen erfolgreichen Übertritt ins Berufsleben betonen.

Kompetenzerwartungen, Interessen, Gewissenhaftigkeit und Passungswahrnehmung als Prädiktoren für Ausbildungsgänge der Sekundarstufe II

Lukas Ramseier, Markus Neuenschwander
University of Applied Sciences Northwestern Switzerland, Schweiz

Theoretischer Hintergrund

Am Ende der Sekundarstufe I können sich Jugendliche in der Schweiz zwischen einer vollzeitschulischen Ausbildung, einer dualen Berufsbildung mit Berufsmaturität und einer dualen Berufsbildung ohne Berufsmaturität entscheiden. Diese drei Ausbildungsarten unterscheiden sich hinsichtlich ihrer Zielperspektive und schulischen Anteile (Baethge, 2006). Weiterhin können Ausbildungsentscheidungen gemäss dem RIASEC-Modell (Holland, 1997) in sechs Berufsfelder

unterteilt werden: ***r**ealistic* (handwerklich-technisch), ***i**nvestigative* (untersuchend-forschend), ***a**rtistic* (kreativ künstlerisch), ***s**ocial* (erziehend-pflegend), ***e**nterprising* (führend-verkaufend) und ***c**onventional* (ordnend-verwaltend; Joerin Fux, 2005).

Die Erwartungs-Wert-Theorie (Eccles & Wigfield, 2002) geht davon aus, dass Erwartungen bezüglich der eigenen Kompetenzen (z.B. Selbstwirksamkeitserwartungen) sowie Werte (z.B. Interessen) Ausbildungsentscheidungen mitbestimmen. Aufgrund bisheriger Forschung wird angenommen, dass darüber hinaus noch weitere Konstrukte relevant sind: Sowohl akademische Selbstkonzepte (Jansen, Scherer & Schroeders) als auch Persönlichkeitsmerkmale wie Gewissenhaftigkeit (Jüttler et al., 2021; Usslepp et al., 2020) weisen Zusammenhänge mit unterschiedlichen Ausbildungsarten auf. Zudem wird auch die subjektiv wahrgenommene Passung zwischen eigenen Bedürfnissen und schulischen Lernangeboten als wichtige Determinante für Ausbildungsentscheidungen angenommen (Nägele, Frey & Neuenschwander, 2017).

Auf Basis dieser Annahmen wird in diesem Beitrag folgende Fragestellung analysiert: Erklären Selbstwirksamkeitserwartungen, Selbstkonzepte, Interessen, Passungswahrnehmung und Gewissenhaftigkeit die Anschlusslösung am Ende der Sekundarstufe I hinsichtlich a) der Ausbildungsart und b) des Berufsfelds?

Methode

Die verwendeten Stichproben stammen aus der dritten sowie vierten Welle des Schweizer Forschungsprojekts [Name der Studie]. Sie bestehen aus 1742 (Modell 1; 49.8% w; 50.2% m) bzw. 1128 (Modell 2; 43.2% w; 56.8% m) Schüler:innen der 7. (Welle 3) bzw. 9. Klasse (Welle 4) aus fünf Schweizer Kantonen.

Zur Überprüfung der Fragestellung wurden zwei multinominale logistische Regressionsmodelle berechnet. In Modell 1 stellt die Ausbildungsart der Anschlusslösung (Berufslehre, Berufsmaturität, Allgemeinbildung) nach der Sekundarstufe I die abhängige Variable dar. Die abhängige Variable in Modell 2 ist das gewählte Berufsfeld (Auswahl der drei häufigsten Berufsfelder: R - handwerklich-technisch, S - erziehend-pflegend und C - ordnend-verwaltend). Beide abhängigen Variablen wurden mit den gleichen Prädiktoren geschätzt: Passungswahrnehmung in der 9. Klasse, Gewissenhaftigkeit, Selbstkonzept Deutsch und Mathematik, berufliche Selbstwirksamkeit und Interesse (operationalisiert durch die sechs Interessenstypen nach Holland, 1997). Als Kontrollvariablen wurden sozioökonomischer Status, Schulniveau, IQ sowie Geschlecht miteinbezogen. Die Variablen wiesen eine befriedigende Reliabilität und Validität auf.

Ergebnisse

Die Regressionsanalysen zeigen, dass die Ausbildungsart (Modell 1) signifikant erklärt wird. Ausbildungsarten mit höherem schulischem Anteil (Berufsmaturität oder Allgemeinbildung) werden durch eine hohe Passungswahrnehmung, ein positives Selbstkonzept im Fach Mathematik sowie ein ausgeprägtes Interesse der I-, A- und E-Dimensionen erklärt. Lediglich die S-Interessendimension weist keinen signifikanten Effekt auf. Ausgeprägte Gewissenhaftigkeit, berufliche Selbstwirksamkeit sowie Interesse in den R- und C-Dimensionen gehen hingegen mit der dualen Berufslehre ohne

Berufsmaturität einher. Der Einfluss des IQ erwies sich nach Einbezug der angenommenen Prädiktoren als nicht signifikant.

Das gewählte Berufsfeld (Modell 2) wird durch Selbstkonzept in Mathematik, berufliche Selbstwirksamkeitserwartung, Gewissenhaftigkeit sowie die R-, I, S- und C-Interessensdimensionen vorausgesagt. Eine Anschlusslösung im ordnend-verwaltenden Bereich geht positiv mit ausgeprägterer Gewissenhaftigkeit sowie Interesse in den S- und C-Dimensionen bzw. negativ mit einem höheren mathematischen Selbstkonzept und Interesse in den R- und I-Dimensionen einher. Eine Anschlusslösung im erziehend-pflegenden Bereich wird durch ausgeprägteres Interesse in den I- und S-Dimensionen vorausgesagt. Das Schulniveau sowie der IQ waren nach Einbezug der weiteren Prädiktoren nicht signifikant.

Bedeutung

Die Ergebnisse zeigen, dass die Erwartungs-Wert-Theorie Ausbildungsentscheidungen sowohl hinsichtlich ihrer Ausbildungsart als auch ihrer fachlichen Ausrichtung vorhersagen. Das Einbeziehen von Persönlichkeitsmerkmalen und Passungswahrnehmung kann dabei die Vorhersagekraft der Erwartungs-Wert-Theorie steigern. Neben Schulniveau, IQ und Geschlecht erklären damit auch individuelle Merkmale Ausbildungsentscheidungen am Ende der Sekundarstufe I. Verschiedene Ausbildungsgänge entsprechen somit unterschiedlichen kognitiven Charakteristiken bzw. Persönlichkeitstypen. Für die pädagogische Praxis ist daher zentral, Jugendliche dabei zu unterstützen, ihre Bedürfnisse und Kompetenzen kennenzulernen, damit sie eine möglichst passende Ausbildung am Ende der Sekundarstufe I wählen können.

A-07: Erwartungs-Wert-Theorien

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S05 T00 B59
Chair der Sitzung: Patrick Paschke

Erfolgserwartungen und Werte als Mediatoren der Effekte von Fähigkeitsselbstkonzept und Selbsteinschätzungsbias auf Mathematiknoten

Patrick Paschke, Ricarda Steinmayr
Technische Universität Dortmund, Germany

[Theoretischer Hintergrund, Fragestellung]

Mehrere Metaanalysen legen nahe, dass hohe Fähigkeitsselbstkonzepte akademischer Kompetenzen in einer bestimmten Domäne (z.B. Mathematik) sich positiv auf den akademischen Erfolg in selbiger Domäne auswirken (Huang, 2011; Valentine et al., 2004; Wu et al., 2021). Aber welche Rolle spielt dabei die Diskrepanz zwischen dem Fähigkeitsselbstkonzept und der tatsächlichen Kompetenz einer Person (*Selbsteinschätzungsbias; SE Bias*)? Ist es für den akademischen Erfolg optimal, seine eigenen Kompetenzen akkurat einzuschätzen, oder ist es vorteilhafter, sie (mäßig) zu über- oder unterschätzen? Oder ist unabhängig von einem SE Bias die Höhe des Fähigkeitsselbstkonzepts an sich bedeutsam für die Leistung von Schüler*innen? Trotz zahlreicher empirischer Studien zu diesem Thema (z.B. Bonneville-Roussy et al., 2017; Lee, 2021), können diese Fragen aufgrund methodischer und statistischer Probleme in bisherigen Studien noch nicht beantwortet werden. Das Kernproblem ist dabei die mangelnde Trennung von Effekten des Fähigkeitsselbstkonzepts und Effekten des SE Bias (Humberg et al., 2019). Somit ist das erste von zwei zentralen Zielen der vorliegenden Untersuchung, die Effekte des Fähigkeitsselbstkonzepts und die Effekte des SE Bias auf akademischen Erfolg getrennt voneinander zu untersuchen.

Sowohl die Effekte des Fähigkeitsselbstkonzepts als auch die Effekte des SE Bias auf den akademischen Erfolg sollen der Theorie nach von motivationalen Variablen mediiert werden (z.B. Bonneville-Roussy et a., 2017; Wigfield & Eccles, 2020). Allerdings wurde diese in Annahme in Bezug auf die Effekte des SE Bias nach unserem Wissen bisher nicht empirisch überprüft. Eines der populärsten Modelle zur Erklärung positiver Fähigkeitsselbstkonzepteffekte auf den akademischen Erfolg ist das Erwartungs-Wert-Modell (Wigfield & Eccles, 2020). Nach diesem Modell wirkt sich das Fähigkeitsselbstkonzept vermittelt über Erfolgserwartungen und Werte (intrinsische Werte, Nützlichkeitswerte und Wichtigkeitswerte) positiv auf den akademischen Erfolg aus. Somit ist das zweite zentrale Ziel der vorliegenden Untersuchung, zu überprüfen, ob die Effekte des Fähigkeitsselbstkonzepts und/oder des SE Bias von Erfolgserwartungen und Werten mediiert werden.

[Methode]

Wir untersuchten 504 Gymnasiat*innen vier zehnter Stufen in NRW zu drei Messzeitpunkten (Alter: M = 15.28; SD = 0.60; 264 weiblich, 240 männlich). Wir erfassten zum ersten und zweiten Messzeitpunkt (*t*1 und *t*2) das Fähigkeitsselbstkonzept in Mathematik, die tatsächliche Kompetenz in Mathematik (mittels eines standardisierten Kompetenztests), sowie die Erfolgserwartung und Werte in Mathematik. Zum zweiten und dritten Messzeitpunkt erfassten wir die Mathematikzeugnisnote als Maß des akademischen Erfolgs in Mathematik. Zudem erfassten wir zum ersten Messzeitpunkt demographische Kontrollvariablen. Die Effekte des *t*1 Fähigkeitsselbstkonzepts, der *t*1 Kompetenz, sowie des *t*1 SE Bias auf die *t*3 Mathematiknote vermittelt über die *t*2 Mediatorvariablen

(Erfolgserwartung und Werte) wurden mit der Response Surface Analysis (Edwards, 2002; Edwards & Parry, 1993; Humberg et al., 2019) untersucht. Hierbei handelt es sich um eine multiple polynomiale Regression, die erlaubt, Effekte des SE Bias und Effekte der Selbsteinschätzung zu trennen und die verschiedenen dargestellten Hypothesen zu SE Bias Effekten gegeneinander zu testen (Humberg et al., 2019). Aufgrund hoher Multikollinearität zwischen den Wertekomponenten überprüften wir einzelne Modelle für jede Komponente.

[Ergebnisse]

In den einzelnen Modellen mediierten die Erfolgserwartungen, die intrinsischen Werte und die Wichtigkeitswerte nicht aber die Nützlichkeitswerte signifikant den Effekt des Fähigkeitsselbstkonzepts auf die Mathenote. Sämtliche Effekte waren positiv und linear. Somit gab es keine quadratischen oder Interaktionseffekte des Fähigkeitsselbstkonzepts und/oder der Kompetenz auf die Mediatoren oder die Mathenote. Diese Befunde sprechen klar gegen SE Bias Effekte auf die Mathenote. Demnach sind Fähigkeitsselbstkonzept und Kompetenz jeweils für sich relevant für den akademischen Erfolg, die Diskrepanz zwischen ihnen hat darüber hinaus aber keinen bedeutsamen Einfluss. Die Ergebnisse werden in Hinblick auf den aktuellen Forschungsstand eingeordnet und kritisch diskutiert.

Wie kann die Motivationspsychologie von verschiedenen Ansätzen der Netzwerkanalyse profitieren?

Miriam Francesca Jähne[1], Alexander Naumann[1], Julia Moeller[2], Julia Dietrich[1]
[1]Friedrich-Schiller-University Jena, Deutschland; [2]Leipzig University, Deutschland

Akademische Motivation ist zentral für den Lernerfolg von Hochschulstudierenden. In den letzten Jahren wurde die Dimensionalität der Facetten der Erwartungs-Wert-Theorie von Eccles und Wigfield (2020) sowohl hinsichtlich ihrer Konzeptualisierung, z.B. dem Stellenwert der Kosten gegenüber den Wertdimensionen (Barron&Hullemann, 2016; Eccles&Wigfield, 2020), als auch hinsichtlich ihrer Messstruktur, z.B. über Bifaktor- und Second-Order-Modelle (Gaspard et al., 2015; Guo et al., 2017), näher in den Blick genommen. Demgegenüber begreift der psychometrische Netzwerkansatz die motivationalen Facetten in einem System komplexer, wechselseitiger Interdependenzen. Diese Blickweise entstammt der Perspektive Dynamischer Systemtheorien (Moeller et al., 2022) und stellt damit innerhalb der Bildungsforschung eine Innovation dar. Erste Studien nutzen bereits netzwerkanalytische Ansätze zur Untersuchung der Leistungsmotivation von Schüler:innen (Tang et al., 2022). Allerdings gibt es bislang nur wenig Anleitung, wie Netzwerkanalysen für die Bildungsforschung nutzbar gemacht werden können.

Unser Beitrag stellt drei Varianten der Netzwerkanalyse vor und verdeutlicht, wie diese (korrelativ sowie partial-korrelativ, Costantini et al., 2018, Co-Endorsement, Moeller et al., 2018) gezielt für die Untersuchung akademischer Motivation eingesetzt und interpretiert werden können. Dabei liegt der Fokus insbesondere auf möglichen Schlussfolgerungen, die aus den jeweiligen Analyseergebnissen abgeleitet werden können und den Fragestellungen, auf die die jeweiligen Analyseansätze antworten. Darüber hinaus werden Empfehlungen hinsichtlich Potentialen und Grenzen der jeweiligen Ansätze gegeben.

Um dieses Ziel zu erreichen, wurden 341 Erstsemesterstudierende (w=70.39%, m=29.33%, d=0.29%; M_{Alter}=20.13(2.44)) unterschiedlicher Fachrichtungen (Biologie, Ernährung, Lehramt, Medizin, Psychologie, Recht, Wirtschaft) querschnittlich am Ende ihres ersten Fachsemesters befragt. Die Rekrutierung der Studierenden (n_{Fach}=39-64) erfolgte vermittelt über ihre Lehrveranstaltungen. In einem Online-Fragebogen wurde erfasst, wie motivierend die Studierenden ihr erstes Fachsemester erlebt haben. Dabei wurden auf Seite der positiven Aufgabenwerte der intrinsische Wert des Studiums, dessen persönliche Wichtigkeit sowie der subjektiv eingeschätzte Nutzen der Studieninhalte erfasst. Aufseiten der negativen Kostenwerte wurden emotionale, Anstrengungs- und Opportunitätskosten eingeschätzt (adaptiert nach Gaspard et al., 2015). Die Erwartungskomponente wurde über das akademische Selbstkonzept operationalisiert (Dickhäuser et al., 2002).

(Partial)korrelative Netzwerke untersuchen die Zusammenhänge zwischen Variablen und ermitteln die Zentralität der Variablen innerhalb des Netzwerks, wobei die partialkorrelativen Netzwerke zusätzlich für den Einfluss aller anderen motivationalen Facetten korrigieren. Die beiden Varianten zielen also auf Aussagen des Typs „Je mehr – desto mehr“. Co-Endorsement-Netzwerke zeigen demgegenüber anhand binärer Daten, inwiefern zwei motivationalen Facetten tatsächlich zugestimmt wird. Diese Unterscheidung ist insofern bedeutsam, als beispielsweise eine positive Korrelation auch dann auftreten kann, wenn zwei Facetten von Motivation eigentlich abwesend bzw. gering ausgeprägt sind.

In den Analysen wiesen sowohl das korrelative als auch das partialkorrelative Netzwerk vorrangig auf Zusammenhänge zwischen den Aufgaben- sowie Kostenwerten untereinander hin. Die Erwartung hatte zudem im partialkorrelativen Netzwerk keine Verbindung zu den übrigen motivationalen Facetten. Hinsichtlich der Zentralität (gemessen durch die Parameter Strength, Closeness und Betweenness, Costantini et al., 2018), d.h. des Einflussreichtums einzelner Wertfacetten innerhalb des Netzwerks, spielten emotionale Kosten (korrelativ) sowie Anstrengungskosten und intrinsischer Wert (partialkorrelativ) jeweils eine wichtige Rolle. Vorläufige Ergebnisse der Co-Endorsement-Netzwerkanalyse verweisen auf eine gehäufte gleichzeitige Zustimmung niedriger emotionaler Kosten mit hohem intrinsischem Wert sowie der Aufgaben- und Kostenwerte untereinander. Den Analysen gemein ist, dass den emotionalen Kosten sowie dem intrinsischen Wert eine relativ prominente Stellung zugewiesen wird.

Je nach angewendeter Netzwerkanalysemethode unterscheidet sich, welche Wertfacetten innerhalb des Netzwerks zentral sind oder besonders häufig auftreten und damit mögliche Ansatzpunkte für Interventionen darstellen. Dabei unterscheidet sich, inwiefern das tatsächliche Auftreten von Wertfacetten oder ein tendenzieller Zusammenhang betrachtet werden soll. Die Wahl des Analyseverfahrens benötigt daher Sorgfalt, um dem Forschungsanliegen gerecht zu werden.

Für zukünftige Untersuchungen der Bildungsforschung ist es zudem wichtig, die motivationalen Netzwerke einzelner Personen oder homogener Personengruppen zu untersuchen. Das Verständnis individueller Profile kann hilfreich sein, um z.B. den Zusammenhang von Studienmotivation und Lernerfolg sowie assoziierter Studienentscheidungen (z.B. Studienabbrüche) besser zu verstehen und passende Unterstützungsangebote zu entwickeln.

Effekte einer Online-Werteintervention auf Selbstkontrolle und Zielerreichung von Studierenden

Melissa Montagna, Oliver Dickhäuser
Universität Mannheim, Germany

Studierende sind tagtäglich mit bewertenden Informationen konfrontiert, die bei ihnen Unsicherheit über die eigene akademische Kompetenz und Zugehörigkeit auslösen und sich negativ auf Wohlbefinden und Leistung auswirken können (Sherman & Cohen, 2006). Gemäß der Self-Affirmation-Theorie erzeugen Informationen, die die eigene Selbstwahrnehmung als grundlegend „integer“ infrage stellen, ein Bedrohungsempfinden und nachfolgend das Bedürfnis, die eigene Selbstintegrität wiederherzustellen (Steele, 1988). Eine Möglichkeit der Wiederherstellung besteht in der Reflexion eigener Werte, die dem Individuum verdeutlicht, dass das Selbst aus mehr als dem bedrohte Teilaspekt besteht (Harackiewicz & Priniski, 2018). Eine Reihe von Studien zeigt, dass das Value-Affirmation-Essay – eine kurze Schreibübung, in der Studierende persönliche Werte reflektieren – beachtliche Effekte in der Leistungsentwicklung von Lerner:innen hervorbringt (Wu et al., 2021). In der vorliegenden Studie versuchten wir diesen Effekt zu replizieren, indem wir den Effekt eine Online-Variante des Value-Affirmation-Essays auf Selbstkontrolle und die akademische Zielerreichung von Studierenden über einen Zeitraum von sechs Wochen betrachteten.

Gemäß der Self-Affirmation Theorie beansprucht der Umgang mit Selbst-bedrohlichen Informationen Kapazität und Motivation (Cohen & Sherman, 2014). Studierende, die aufgrund der Interventionsübung effektiver mit selbstbedrohlichen Informationen umgehen können, sollten demnach über mehr Kapazität für die Bewältigung anderer Anforderungen im universitären Kontext verfügen. Aus diesem Grund nahmen wir an, dass sich die Intervention positiv auf den Selbstkontrolleinsatz von Studierenden auswirken sollte. Eine Vielzahl vorhergehender Value-Affirmation-Studien berichtet zudem von einer primären Wirksamkeit der Intervention für Studierende aus Minderheitengruppen (z.B. African-American-Studierende, First-Generation-Studierende; für einen Überblick siehe Wu et al., 2021). Der Grund für die erhöhte Wirksamkeit der Intervention liege in der chronisch erhöhten Konfrontation mit Selbstintegritäts-bedrohenden Informationen, denen diese Minderheitsgruppen im akademischen Kontext ausgesetzt seien (Cohen et al., 2006). Vorhergehende Untersuchungen haben jedoch nicht zwischen der subjektiven Bedrohungserfahrung und demographischen Gruppenindikatoren differenziert (Easterbrook et al., 2021). Aus diesem Grund wurde in der vorliegenden Studie untersucht, ob ein potenziell stärker ausgeprägter Interventionseffekt für Studierende mit Migrationshintergrund oder geringerem familiären Bildungshintergrund durch ein höheres Bedrohungsempfinden zu erklären ist.

Es wurde ein gemischtes längsschnittliches experimentelles Design implementiert mit wahrgenommener Bedrohung und Interventionsbedingung als Between-Subjects-Faktoren und Zeit als Within-Subjects-Faktor. Der Effekt der 15-minütigen Online-Intervention auf Selbstkontrolleinsatz und Zielerreichung wurde über einen Zeitraum von sechs Wochen untersucht (N = 127 Studierende, 694 Beobachtungen).

Das Value-Affirmation-Essay erhöhte Selbstkontrolleinsatz, $b = 0.21$, $t(127) = 2.57$, $p = .01$, und Zielerreichung, $b = 0.50$, $t(126) = 2.89$, $p < .01$. Zudem zeigte sich, dass der Effekt auf Zielerreichung vollständig über einen erhöhten Selbstkontrolleinsatz vermittelt wurde, $b = 0.80$, $p < .001$. In

explorativen Analysen zeigte sich zudem eine signifikante Interaktion der Interventionsbedingung mit der Zeit: Teilnehmer:innen, die die Value-Affirmation-Übung erhalten hatten, zeigten einen sich über die Zeit erhöhenden Effekt der Intervention auf Selbstkontrolleinsatz, $b = 0.05$, $t(579) = 2.22$, $p = .03$ und Zielerreichung, $b = 0.08$, $t(575) = 2.03$, $p = .04$. Entgegen den Erwartungen zeigte sich kein differenzieller Effekt der Intervention auf den Selbstkontrolleinsatz in Abhängigkeit demographischer Gruppenindikatoren (hier: Migrationshintergrund, familiärer Bildungshintergrund), ps > .39 sowie keine Moderation des Interventionseffekts durch ein erhöhtes Bedrohungsempfinden, $p = .13$.

Zusammenfassend liefert der Beitrag experimentelle Evidenz für die Annahme, dass eine kurze Online-Intervention zur Reflexion eigener Werte positive Effekte auf den Selbstkontrolleinsatz und die Zielerreichung von Studierenden hat, die über die Zeit zunehmen. Entgegen theoretischen Erwartungen ging ein erhöhtes Bedrohungsempfinden jedoch nicht mit einer erhöhten Wirksamkeit der Intervention einher. Dieses Ergebnis weist möglicherweise auf die Notwendigkeit hin, in nachfolgenden Untersuchungen auf die dynamische Prozess-Beziehung zwischen intraindividuellen Erfahrungen von Bedrohungen, darauffolgenden Reaktionen und den intervenierenden Einfluss der Value-Affirmation-Intervention zu fokussieren. Die geringen Kosten der Interventionsimplementation (online, geringer Zeitaufwand) sowie die Wirksamkeit unabhängig von demographischen Vulnerabilitätsindikatoren und die wachsenden Interventionseffekte unterstreichen die Wirksamkeit von Value-Affirmation-Interventionen und sprechen für eine frühe und breite Anwendung in Bildungskontexten.

Erwartungs-Wert-Konstrukte und Basic Needs als Prädiktoren für die Leistung im datenbasierten mathematischen Modellieren

Matthias Mohr, Stefan Ufer
LMU München, Germany

Theoretischer Hintergrund

Lehr-Lern-Labore zielen darauf ab, Lernenden einen außerschulischen Zugang zu zentralen fachlichen Arbeitsweisen zu ermöglichen, wie beispielsweise dem Modellieren mit Daten in der Mathematik. Es ist zu erwarten, dass sowohl die dispositionalen Überzeugungen der Person zum Inhalt, als auch das konkrete Erleben des Lernsettings die Ergebnisse der Arbeit im Labor beeinflussen.

Die Erwartungs-Wert Theorie (EWT) beschreibt den Einfluss von Erwartungs- und Wertüberzeugungen sowie von Kosten auf die Leistung von Schüler:innen (Eccles & Wigfield, 2020). Obwohl der situative Charakter dieser Überzeugungen in letzter Zeit stärker betont wird, spiegeln die Überzeugungen der EWT vor allem dispositionale Einstellungen einer Person wider. Aussagen über das situative Erleben einer Lernsituation sind im Rahmen der EWT nur beschränkt möglich. Die Selbstbestimmungstheorie (SDT, Ryan & Deci, 2020) setzt hier an und postuliert die Erfüllung psychologischer Grundbedürfnisse (Basic Needs) als Voraussetzungen für ein positives situationales Erleben. Grundsätzlich ist anzunehmen, dass die Erfüllung der Basic Needs auch von den initialen Überzeugungen der Lernenden abhängt. So dürfte z.B. das Kompetenzerleben auch von den initialen Einschätzungen der eigenen Fähigkeiten beeinflusst werden. Gleichermaßen gehen SDT und SEVT

davon aus, dass ihre Konstrukte wesentliche Voraussetzungen für situationale Motivation und Anstrengung beschreiben (u.a. Dietrich et al., 2017; Goldman et al., 2017).

Fragestellung

Ziel dieser Studie war es den Einfluss von aufgabenspezifischen Erwartungs- und Wertüberzeugungen und die Erfüllung der Basic Needs auf die Leistung zum datenbasierten Modellieren in einem Lehr-Lern-Labor zu untersuchen. Es wurde angenommen, dass die Basic Needs den Einfluss der Erwartungs- und Wertüberzeugungen sowie der Kosten auf die situationale Motivation und die Anstrengung vermitteln (Hypothese 1) und dass alle genannten Faktoren dann direkt bzw. indirekt zu Erklärung interindividueller Unterschiede in der Leistung beitragen (Hypothese 2).

Methode

Im Rahmen eines Lehr-Lern-Labors erarbeiteten sich *N=369* Schüler:innen anhand exemplarischer Problemstellungen die Prinzipien datenbasierten Modellierens. Anschließend wendeten sie die in einer weiteren Aufgabe selbstständig an. Deren Bearbeitung wurde zur Messung der Leistung herangezogen. Direkt vor dieser Anwendungsphase wurden die Schüler:innen zu ihren aufgabenspezifischen Erwartungs- (α=.91) und Wertüberzeugungen (α=.86) sowie zu den Kosten (α=.79) befragt. Nach der Anwendungsphase wurden die Erfüllung der Basic Needs (Gesamtskala, α=.69), die situationale Motivation (extrinsisch bis intrinsisch, α=.75-.90) und die Anstrengung (α=.86) erhoben. Im Rahmen eines Tests vor der Bearbeitung der Lernumgebung wurde das Vorwissen zu Funktionen (α=.72) gemessen. Es wurden Strukturgleichungsanalysen in *R* mit dem package „*lavaan*“ durchgeführt.

Ergebnisse

Wie erwartet sagen die Erwartungs- ($\beta=.34, p<.001$) und Wertüberzeugungen ($\beta=.28, p<.001$) sowie die Kosten ($\beta=-.23, p<.001$) das Erleben der Basic Needs ($R^2=.48$) vorher. Ebenfalls, wie erwartet, sind die Basic Needs (Mot.: $\beta=.48, p<.001$; *Anst.:* $\beta=.53, p<.001$) prädiktiv für die situationale Motivation ($R^2=.73$) und die Anstrengung ($R^2=.43$), wobei ein signifikanter direkter Effekt der Wertüberzeugungen (Mot.: $\beta=.41, p<.001$; *Anst.:* $\beta=.15, p<.05$) verbleibt. Die Leistung ($R^2=.55$) wird durch die Anstrengung ($\beta=.36, p<.01$) und durch das Vorwissen ($\beta=.50, p<.001$), jedoch nicht durch die situationale Motivation ($\beta=.23, p=.167$) erklärt. Es gibt keine weiteren signifikanten direkten Pfade. Die indirekten Effekte mit Mediator Basic Needs bzw. mit Mediatoren Anstrengung oder Motivation waren signifikant.

Diskussion

Die Ergebnisse bestätigen die angenommenen Zusammenhänge zwischen den beiden Theoriebereichen. Sie untermauern die Rolle der Basic Needs als Initiator situationaler Motivation und Anstrengung, weisen aber auch darauf hin, dass dieses situationale Erleben nicht nur von der Situation, sondern auch von Erwartungs- und Wertüberzeugungen bzw. Kosten beeinflusst wird. Die Basic Needs vermitteln zwischen diesen individuellen Überzeugungen und der situationalen Motivation bzw. der Anstrengung. Der verbleibende direkte Effekt der Wertüberzeugungen weist

darauf hin, dass weitere vermittelnde Mechanismen anzunehmen sind (z.B. über Emotionen, Pekrun et al., 2007). Auch wenn die Erwartungs- und Wertüberzeugungen sowie Kosten nur indirekt auf die Leistung wirken, stellen sie doch bedeutsame Faktoren zur Initiierung eines positiven situationalen Erlebens dar. Es werden also die Konstrukte beider Theoriebereiche benötigt, um Unterschiede in der Leistung zu erklären.

A-08: Heterogenität Grundschule

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S05 T00 B71
Chair der Sitzung: Svenja Maria Hascher

Heterogenität in der Grundschulklasse und ihre Auswirkungen auf Leistung. Welche Rolle spielt die Unterrichtsqualität?

Svenja Maria Hascher, Camilla Rjosk, Georg Lorenz
Institut zur Qualitätsentwicklung im Bildungswesen, Germany

Theorie und Forschungsstand

Die (steigende) Heterogenität der Schülerschaft ist ein wichtiges Thema des bildungswissenschaftlichen Diskurses (z.B. Vock und Gronostaj, 2017). Zu der Frage, wie sich die Heterogenität in Grundschulklassen auf die Leistung der Schüler:innen auswirkt, existieren unterschiedliche Theorien und Befunde. Eine zentrale Annahme des vorliegenden Beitrags ist, dass die Unterrichtsqualität einen Einfluss auf den Zusammenhang von Heterogenität und Leistung hat.

Heterogenität kann potenziell Lernprozesse und damit Leistung positiv beeinflussen, indem die Schüler:innen voneinander lernen und durch die Begegnung mit unterschiedlichen Hintergründen ihre Kompetenzen ausbauen. Dies geschieht, indem die kognitive Dissonanz, die sich durch die Begegnung mit Unbekanntem ergibt, aufgelöst wird und so eigene Denkprozesse restrukturiert und weiterentwickelt werden (Palincsar 1998; Vygotsky 1978; Piaget 1977). Teilweise bestätigten bisherige Studien positive Zusammenhänge zwischen verschiedenen Heterogenitätsdimensionen und schulischen Leistungen (Scharenberg 2014; Dotzel et al. 2021; Van der Slik et al. 2006). Aufbauend auf Theorie (Lipowsky 2020) und Forschungsergebnissen (van Leeuwen und Janssen 2019) zu Gruppenlernprozessen wird angenommen, dass Lehrkräfte, die in den drei tiefenstrukturellen Merkmalen von Unterrichtsqualität (kognitive Aktivierung, unterstützendes Unterrichtsverhalten, Klassenmanagement) (Klieme et al. 2001) hohe Werte erreichen, positive Effekte von Heterogenität verstärken können, indem sie soziales Lernen positiv beeinflussen und kognitive Dissonanz nutzbar machen.

Heterogenität kann jedoch potenziell auch negative Effekte auf Leistung haben. Diese können sich durch sog. „intergroup conflict“ ergeben (Social Identity Theory; Tajfel und Turner 1979). Auch zu negativen Effekten von Heterogenität liegen Befunde vor (z.B. Marotta 2017; Conway-Turner et al. 2020). Lehrkräfte könnten negative Auswirkungen von Heterogenität abmindern, indem sie durch unterstützendes Unterrichtsverhalten Konflikte vermindern (Hendrickx et al. 2016) und durch Klassenmanagement einen positiven Einfluss auf das Verhalten der Schüler:innen ausüben (Korpershok et al. 2016).

Der eigene Beitrag knüpft an Befunde an, die eine besondere Bedeutung von kognitiver Aktivierung und unterstützendem Unterrichtsverhalten in leistungsheterogenen Klassen bestätigen (Decristan et al. 2017).

Fragestellungen

1) Wie wirken sich verschiedene Dimensionen klasseninterner Heterogenität auf die Leistung der Schüler:innen aus?

2) Wie beeinflusst Unterrichtsqualität die Zusammenhänge zwischen Heterogenität und Leistung?

Methode

Es werden Sekundäranalysen von Daten des Nationalen Bildungspanels (NEPS; vgl. Blossfeld & Roßbach, 2019) zu drei Messzeitpunkten durchgeführt (1881 Schüler:innen in 239 2. Klassen, 238 Mathematiklehrkräfte). Die Klassenheterogenität wurde in der zweiten Klasse erhoben (t1) (Leistungsheterogenität und sozio-ökonomische Heterogenität gemessen anhand der Standardabweichung, ethnische Heterogenität gemessen mit dem Diversitätsindex Simpson's D). Die Lehrkräftedaten zur Unterrichtsqualität wurden in Klasse 3 (t2) erfasst (Lehrkräftefragebogen, zwei Items pro Unterrichtsqualitätsmerkmal). Die abhängige Variable ist die Mathematikleistung in der vierten Klasse (t3). In Mehrebenenanalysen werden zunächst Auswirkungen von den Heterogenitätsdimensionen auf die Mathematikleistung analysiert, bevor im zweiten Schritt Interaktionsterme mit den drei Merkmalen von Unterrichtsqualität hinzugenommen werden, um mögliche Moderationseffekte aufzudecken. Kontrollvariablen sind mittlere Leistung und sozio-ökonomischer Hintergrund sowie Schüler:innenanteil mit Migrationshintergrund zu t1.

Ergebnisse

Die folgenden vorläufigen Ergebnisse beruhen auf Regressionsanalysen auf Klassenebene.

Zunächst wurden kleine positive Effekte von ethnischer und sozio-ökonomischer Heterogenität auf Leistung gefunden. Interaktionseffekte bestätigen die Annahme der Moderation der Auswirkungen verschiedener Heterogenitätsdimensionen durch kognitive Aktivierung und unterstützendes Unterrichtsverhalten. Eine hohe kognitive Aktivierung sorgt für einen positiven Einfluss von Leistungsheterogenität und sozio-ökonomischer Heterogenität auf Leistung.

Für das unterstützende Unterrichtsverhalten findet sich die Leistungsheterogenität betreffend ein kontraintuitiver Befund: Ein unterstützendes Unterrichtsverhalten scheint dazu zu führen, dass sich Leistungsheterogenität negativ auf die Leistung auswirkt.

Ein erster Erklärungsansatz könnte sein, dass eine Lehrkraft, die sich vornehmlich darauf konzentriert, die Schüler:innen emotional zu unterstützen, weniger Aufmerksamkeit auf die kognitive Aktivierung legt, die sehr bedeutsam für die positiven Effekte von Leistungsheterogenität ist.

Der kleine positive Einfluss der ethnischen Heterogenität scheint durch Unterrichtsqualität nicht beeinflusst zu werden.

Vor dem Hintergrund dieser Ergebnisse und weiterer Analysen wird die besondere Bedeutung von Unterrichtsqualität in heterogenen Gruppen diskutiert. Weiterführend sind differenzielle Mehrebenenanalysen zu Schüler:innen mit unterschiedlichen Hintergrundmerkmalen geplant.

Der Schuleintritt von Kindern mit Beeinträchtigung: Ein systematisches Review internationaler Befunde

Daniel Then, Sanna Pohlmann-Rother
Julius-Maximilians-Universität Würzburg, Deutschland

Der Schuleintritt ist eine zentrale Statuspassage im Leben eines Kindes und seiner Familie (Griebel & Niesel, 2020). Kinder mit Beeinträchtigung und deren Eltern sind dabei mit spezifischen Herausforderungen konfrontiert: Die Kinder sind gefordert, mit Änderungen in den sie umgebenden Unterstützungsstrukturen umzugehen, welche nur selten nahtlos vom vorschulischen an den

schulischen Bereich anschließen (Daley et al., 2011). Die Eltern stehen vor der Aufgabe, zusätzliche Unterstützung für ihr Kind zu organisieren (Janus & Siddiqua, 2018) und die passende Schulform (Regel- vs. Förderschule) zu wählen (Dorrance, 2010). Um Kinder und Eltern im Übergangsprozess adäquat zu unterstützen, gilt aus ökosystemischer Perspektive der Einbezug aller am Übergang beteiligten AkteurInnen als grundlegend (Rimm-Kaufman & Pianta, 2000). Die internationale Forschungslage zur Frage, in welchem Umfang und mit welchen Zielrichtungen die AkteurInnen im Kontext eines inklusiven Übergangs interagieren und welche Herausforderungen oder Chancen sich beim Schuleintritt von Kindern mit Beeinträchtigung identifizieren lassen, ist schwer zu überblicken und wurde bislang nicht systematisch aufgearbeitet. So fehlt es an einem systematischen Überblick über Forschungsarbeiten, welche den Schuleintritt von Kindern mit Beeinträchtigung fokussieren und nach Verabschiedung der UN-Behindertenrechtskonvention 2006 durchgeführt wurden. Dies erschwert es, Bedarfe für weitere Forschung in diesem Bereich zu identifizieren und Evidenz zu generieren, auf deren Grundlage Maßnahmen zur inklusiven Gestaltung des Schuleintritts für Kinder mit Beeinträchtigung abgeleitet werden können. Im Rahmen des Vortrags wird diesem Desiderat nachgegangen, indem folgende Fragestellungen thematisiert werden:

1. Auf welche Stichproben und Untersuchungsdesigns wird in Studien zum Schuleintritt von Kindern mit Beeinträchtigung zurückgegriffen, die nach Verabschiedung der UN-Behindertenrechtskonvention 2006 durchgeführt wurden?
2. Welche Schwerpunkte lassen sich in der Forschung zum Schuleintritt von Kindern mit Beeinträchtigung identifizieren?
3. Welche Forschungsdesiderate bestehen im Hinblick auf den Schuleintritt von Kindern mit Beeinträchtigung?

Das methodische Vorgehen orientierte sich am PRISMA-Flow-Diagramm (Page et al., 2021). Zunächst wurden in einer systematischen Datenbankrecherche sechs wissenschaftliche Datenbanken (PsycInfo, Web of Science, ERIC, JSTOR, BIDOK, Fachportal Pädagogik) durchsucht. Aus einem Korpus von 4.560 Studien wurden in einem mehrschrittigen Vorgehen 31 Studien identifiziert, die ins Review aufgenommen wurden. Eine Suche in zusätzlichen Quellen (Google Scholar, Publikationslisten, Literaturverzeichnisse, thematisch verwandte Reviews) ergab weitere 774 Studien, aus denen in einem mehrschrittigen Selektionsverfahren weitere 24 Studien für die Aufnahme ins Review ermittelt wurden. Die Gesamtzahl der aufgenommenen Studien ergab somit 55. Sowohl bei der systematischen Datenbankrecherche als auch bei der Durchsicht zusätzlicher Quellen erfolgte die Selektion der Studien durch zwei unabhängige Forschende, die die Auswahl der Studien diskutierten und eine Einigung erzielten. Die Auswertung erfolgte deduktiv mithilfe eines Kategoriensystems, welches aus einem von den AutorInnen entwickelten, ökosystemischen Übergangsmodell abgeleitet wurde. Die Zuordnung der Studien zu den Kategorien wurde von zwei unabhängigen CodiererInnen vorgenommen (*Cohens Kappa*=.80).

Die Auswertung der Studien zeigt, dass in empirischen Erhebungen zum Schuleintritt von Kindern mit Beeinträchtigung vielfältige Methoden genutzt werden, wenngleich (qualitative) Interview- und (quantitative) Fragebogenerhebungen dominieren. Ein starker Fokus der Studien liegt dabei auf Kindern mit entwicklungsbezogen Beeinträchtigungen (v.a. Autismus). Ein inhaltlicher Schwerpunkt bisheriger Forschungsarbeiten liegt auf den subjektiven Perspektiven der Beteiligten, insbesondere der Eltern, PädagogInnen und MitarbeiterInnen zusätzlicher Unterstützungssysteme (z.B. ErgotherapeutInnen) die in das Übergangsgeschehen eingebunden sind. Im Mittelpunkt stehen

wahrgenommene Barrieren (z.B. mangelnde Expertise der Fach- und Lehrkräfte) und Unterstützungsfaktoren im Übergang (z.B. hinreichende Expertise der MitarbeiterInnen zusätzlicher Unterstützungssysteme). Die Prozesse, welche den erfolgreichen Verlauf des Übergangs moderieren (z.B. die adaptive Förderung des Kindes), waren bisher dagegen eher selten primärer Betrachtungsgegenstand. Auch Studien, welche die Strukturen des Bildungssystems bzw. den administrativen Rahmen des Übergangsprozesses in den Blick nehmen (z.B. Zurückstellungen und ihre Auswirkungen auf Kinder mit Beeinträchtigung), sind bislang selten. Der vorliegende Beitrag gibt somit Hinweise auf Forschungslücken, die zu füllen bedeutende Voraussetzung ist, um den Übergangsprozess für Kinder mit Beeinträchtigung entwicklungsförderlich gestalten und evidenzbasiert weiterentwickeln zu können.

Inklusives pädagogisches Wissen von Studierenden des Grundschullehramts zwischen Unsicherheit und Evidenz

Saskia Liebner[1], Christian Elting[2], Sonja Ertl[1]
[1]Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland; [2]Otto-Friedrich-Universität Bamberg

Die Grundschule ist als grundlegende und gemeinsame Schule für (fast) alle Kinder konzipiert und somit auf Inklusion angelegt. Dies erfordert eine entsprechende Professionalisierung von Grundschullehrkräften für einen inklusiven Umgang mit Heterogenität. Empirisch gestützte Modelle der professionellen Kompetenz (z.B. Baumert & Kunter, 2011) verweisen auf die Bedeutung entsprechender Überzeugungen und Wissensbestände. Inklusives pädagogisches Wissen, das neben allgemeinen Kompetenzen im Bereich Diagnose, Intervention etc. insbesondere spezifisches Wissen zur Inklusion bzw. zu inklusivem Unterricht umfasst (König et al., 2019), ist hierbei zentral, aber im Grundschullehramtsstudium weitgehend unerforscht. Zu inklusiven Überzeugungen liegen zwar Befunde vor (z.B. Strauß & König, 2017) – Studienanfänger*innen werden dabei aber kaum berücksichtigt. Auch sind Zusammenhänge zwischen Merkmalen des Grundschullehramtsstudiums, dem Erwerb inklusiven pädagogischen Wissens und inklusiven Überzeugungen ungeklärt. Diesen Desiderata widmet sich die vorliegende Studie. Dabei wird ein weites Verständnis von Inklusion zugrunde gelegt, das nicht nur Kinder mit sonderpädagogischem Förderbedarf berücksichtigt, sondern *alle* Kinder mit besonderen pädagogischen Bedürfnissen (Elting et al., 2019).

Der Beitrag geht folgenden Teilfragestellungen nach:

(1) Unterscheiden sich beginnende und fortgeschrittene Grundschullehramtsstudierende in ihren inklusiven Überzeugungen und ihrem inklusiven pädagogischen Wissen?

(2) Lassen sich inklusive Überzeugungen durch Merkmale des Studiums oder das inklusive pädagogische Wissen der Grundschullehramtsstudierenden erklären?

In einer Querschnittserhebung wurden N = 716 ($n1$ = 552 beginnende; $n2$ = 164 fortgeschrittene) Grundschullehramtsstudierende zu ihrem inklusiven pädagogischen Wissen und ihren inklusiven Überzeugungen befragt. Das inklusive Wissen wurde über eine Adaption der Instrumente GPK-IT (Gerhard et al., 2020) und BilWiss-2.0 (Kunina-Habenicht et al., 2020) erfasst. Die Reliabilität der Gesamtskala (30 Items; Kodierung: 0 bis 2 Punkte pro Item; α = .75) ist für inhaltlich breit angelegte Wissenstests zufriedenstellend. Zur Erfassung der inklusiven Überzeugungen kamen KIESEL-

Markerskalen (Bosse & Spörer, 2014; 4 Items pro Skala; Rating: 1 = lehne voll ab, 4 = stimme voll zu; Einstellung: α = .62; Selbstwirksamkeit: α = .73) zum Einsatz.

Inferenzstatistische Analysen zu Forschungsfrage 1 zeigen, dass fortgeschrittene im Vergleich zu beginnenden Grundschullehramtsstudierenden etwas pro-inklusiver eingestellt ($t(714) = 1.59$, $p < .10$, $d = 0.14$) und etwas weniger selbstwirksam sind ($t(714) = -3.98$, $p < .001$, $d = -0.35$) und dabei über ein deutlich umfangreicheres inklusives pädagogisches Wissen verfügen ($t(714) = 10.40$, $p < .001$, $d = 0.93$). Die mittleren Überzeugungs- und Wissenswerte liegen jedoch für beide Gruppen in einem positiven Bereich, signifikant oberhalb der theoretischen Mitte ($p < .001$).

Mit Blick auf Forschungsfrage 2 zeigen multiple lineare Regressionsanalysen unter Kontrolle von Hintergrundmerkmalen der Person, dass Merkmale des Studiums einen kleinen Beitrag zur Erklärung inklusiver Überzeugungen leisten (für Einstellung und Selbstwirksamkeit jeweils korrigiertes R^2 = .19). Die Höhe des Fachsemesters ist dabei nicht signifikant mit inklusiven Einstellungen verknüpft und negativ mit inklusiven Selbstwirksamkeitserwartungen ($\beta = -.20$, $p < .01$). Der Besuch inklusiver Lehrangebote, inklusiver Praktika und die Belegung der Didaktik des Deutschen als Zweitsprache sind dementgegen mit etwas pro-inklusiveren Einstellungen ($\beta = .13$, $p < .01$; $\beta = .06$, $p < .10$; $\beta = .07$, $p < .10$) und höheren Selbstwirksamkeitserwartungen ($\beta = .09$, $p < .05$; $\beta = .10$, $p < .05$; $\beta = .07$, $p < .05$) assoziiert. Ein umfangreicheres inklusives pädagogisches Wissen hängt mit etwas pro-inklusiveren Einstellungen ($\beta = .16$, $p < .01$) und weniger stark ausgeprägten inklusiven Selbstwirksamkeitserwartungen ($\beta = -.08$, $p < .05$) zusammen.

Die in ihrer Reichweite limitierten Querschnittsbefunde liefern erste empirische Belege dafür, dass weder das Fortschreiten im Studium noch die Vermittlung inklusiven pädagogischen Wissens per se mit pro-inklusiveren Überzeugungen gleichzusetzen sind. Offenbar bedarf es einschlägiger Lehrangebote. Die geringere Selbstwirksamkeit der fortgeschrittenen Studierenden könnte auf Unsicherheiten im Umgang mit bevorstehenden Herausforderungen zurückzuführen sein. Die Befunde werden mit Blick auf die Professionalisierung angehender Grundschullehrkräfte für einen inklusiven Umgang mit Heterogenität diskutiert und weitere Forschungsdesiderata aufgezeigt.

A-09: Schulleitung und Schulaufsicht

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S05 R03 H20
Chair der Sitzung: Fatmana Selcik

Schulleitung im Spannungsfeld zwischen Heterogenität und Disruption

Fatmana Selcik[1], Katharina Reschke[1], Nina Jude[1], Stefan Brauckmann-Sajkiewicz[2], Tobias Feldhoff[3], Katharina Maag Merki[4], Falk Radisch[5], Francesca Suter[6]

[1]Universität Heidelberg; [2]Alpen-Adria-Universität Klagenfurt, Österreich; [3]Johannes Gutenberg-Universität Mainz; [4]Universität Zürich, Schweiz; [5]Universität Rostock; [6]Pädagogische Hochschule Graubünden, Schweiz

Theoretischer Hintergrund

Im Sinne des Konzeptes der Einzelschule als pädagogische Handlungseinheit (Fend, 1986) können Herausforderungen im Zusammenhang mit Heterogenität in Schulen unterschiedlich wahrgenommen werden und damit auch unterschiedlich auf die Qualität und Weiterentwicklung von Schule und Unterricht (van Ackeren et al., 2021) sowie auf die Förderung der Schüler:innen wirken (Maag Merki et al., 2021). Auch das Schulleitungshandeln ist durch Kontextbedingungen der Schule geprägt (Klein, 2016). Schulleitungshandeln ist dabei eingebettet in die sich gegenseitig bedingenden Prozesse der Organisationsentwicklung (OE), Personalentwicklung (PE) und Unterrichtsentwicklung (UE) (Rolff, 2019). In diesem Gefüge gewinnt Führungshandeln an Bedeutung (Leithwood & Louis, 2012). Effektive Führung kann als eine der wichtigsten Ressourcen für Organisationen sowohl in Zeiten der Normalität als auch in Krisenzeiten angesehen werden (Shufutinsky et al., 2020). Die Perspektive von Schulleitenden ist daher zentral für die erfolgreiche Steuerung des Lehrens und Lernens (Altrichter et al., 2019; Helm et al., 2016). Die durch die Corona-Pandemie ausgelösten Disruptionen haben Schulleitende besonders gefordert, schulische Bildung aufrechtzuerhalten und schnell und vorausschauend zu handeln (Harris & Jones, 2020).

Fragestellung

Mit dem Blick auf die Einzelschulen und die Heterogenität der jeweiligen Kontextbedingungen wird angenommen, dass Schulen die Pandemiebewältigung unterschiedlich wahrgenommen haben. Wir gehen daher folgender Fragestellung nach: Lassen sich *systematische Unterschiede zwischen Schulen* zu Beginn der Pandemie hinsichtlich ihrer Ressourcen und der wahrgenommenen Herausforderungen aus Sicht der Schulleiter:innen identifizieren? Wir erwarten, dass (H1) Schulen ihre *Entwicklungsprozesse je nach Kontextbedingungen unterschiedlich gestaltet* haben und in den Bereichen OE, PE, UE überdurchschnittliche oder unterdurchschnittliche Werte aufweisen. Zusätzlich erwarten wir (H2) eine *unterschiedliche Entwicklung der Schulen über die Zeit* in diesen Bereichen.

Methode

Die Analysen basieren auf Daten der tri-nationalen Längsschnittstudie „S-CLEVER. Schulentwicklung vor neuen Herausforderungen“. Diese untersucht, wie Schulen mit den Herausforderungen im ersten Jahr der Pandemie umgegangen sind und inwiefern bzw. in welchen Bereichen sie ihre Schule systematisch weiterentwickelt haben. Schulleitende der allgemeinbildenden Schulen in Deutschland, Österreich und der Deutschschweiz wurden zu drei Zeitpunkten im Verlauf des Schuljahres 2020/2021 befragt. Anhand latenter Profilanalysen (LPA) wurden Subgruppen von Schulen hinsichtlich ihrer Erfahrungen mit digitalem Lernen, den wahrgenommenen Herausforderungen sowie die wahrgenommene Belastung und Zufriedenheit während der Pandemie identifiziert. Auf der Ebene des Unterrichts (UE) wurden die Sicherstellung der Unterrichtsqualität und die Weiterentwicklung des

Unterrichts, auf der Ebene der Organisation (OE) der Unterstützungsbedarf in Bezug auf die Digitalisierung und auf der Personalebene (PE) die wahrgenommene Professionalisierung erfasst.

Ergebnisse und ihre Bedeutung

Mittels der LPA konnten drei Profile unterschieden werden. Schulen in Profil 1 wiesen bei den profilbildenden Variablen eine hohe Erfahrung mit digitalem Lernen auf und berichteten geringe Herausforderungen zu Beginn des Schuljahres. Allerdings berichteten sie im Verlauf des Schuljahres einen signifikanten Anstieg ihrer Belastungen. Schulen in Profil 2 berichteten von geringen Erfahrungen mit digitalem Lernen und sehr hohen Herausforderungen zu Beginn der Pandemie. Schulen in Profil 2 haben den größten Anteil an Heterogenität in der Schülerschaft und starteten mit vielfältigen Herausforderungen personeller sowie technischer Art. Schulen in Profil 3 wiesen neben den höchsten Erfahrungen in Bezug auf digitales Lernen mittlere Werte auf den anderen Variablen auf.

Auf der Ebene der Organisation berichteten die Schulen in Profil 2 im Verlauf des Schuljahres signifikant weniger Unterstützungsbedarf. Hinsichtlich der Möglichkeit individueller Lernförderung durch Digitalisierung scheinen hier Chancen erkannt und umgesetzt worden zu sein. Auf der Ebene des Personals gab es zu Beginn des Schuljahres Unterschiede in der wahrgenommenen Professionalisierung. Während Schulen in Profil 1 von einer signifikant hohen wahrgenommenen Professionalisierung berichteten, erlebten sich Schulen in Profil 2 weniger professionell. Schulen in Profil 1 und 2 zeigen Veränderungen auf den Ebenen Organisation und Personal. Der Beitrag diskutiert Implikationen für die Praxis von Schulleitungshandeln in heterogenen Spannungsfeldern.

Die Rolle der Schulaufsicht in der Implementation von Ganztag – Ergebnisse einer Interviewstudie

Nicole Zaruba[1], Raphaela Porsch[1], Larissa Habeck[2], Falk Radisch[2]
[1]Otto-von-Guericke-Universität Magdeburg, Germany; [2]Universität Rostock

Ganztagsschulen wird eine zentrale Rolle zugeschrieben, wenn darum geht, die Bildungsgerechtigkeit zu erhöhen und den heterogenen Voraussetzungen der Schülerschaft gerecht zu werden (KMK, 2002). In dem komplexen Schulentwicklungsprozess zur Ganztagsschule sind verschiedene Akteur:innen involviert, deren Überzeugungen und Handeln wesentlich zum Gelingen dieser Schulform beitragen können – so auch die Schulaufsicht. Welche Überzeugungen die Akteur:innengruppe der Schulaufsicht zur Reduktion von Bildungsbenachteiligung hat, ist nicht erforscht. In der Befragung von Buchna et al. (2017) spielte die Reduktion von Bildungsbenachteiligung für z. B. Schulleitungen und pädagogisches Personal explizit keine Rolle.

Die Schulaufsicht hat neben ihrer genuinen Aufsichtsrolle die Aufgabe, Schulen durch Beratung zu unterstützen (Dedering, 2021). Obwohl Beratung seitens der Schulaufsicht in vielen Schulgesetzen für Schulentwicklungsprozesse, insbesondere für die Entwicklung zur Ganztagsschule, explizit gefordert wird, wurden Mitarbeitende der Schulaufsicht dazu bislang nicht untersucht. Da Überzeugungen als Handlungs- und Wahrnehmungsfilter wirken können (Kunter & Pohlmann, 2015), ist es plausibel, davon auszugehen, dass die Überzeugungen zum Ganztag seitens der Schulaufsichtsbeamt*innen ihr berufliches Handeln beeinflussen und sie somit einen wesentlichen Beitrag zum Gelingen dieser Bildungsreform leisten können.

Im Rahmen des Vortrags werden folgende Fragestellungen adressiert:

1. Welche Überzeugungen haben Mitarbeitende der Schulaufsicht zum Ganztag?
2. Bei welchen Akteur*innen sehen Schulrät*innen die Verantwortung in der Entwicklung zur Ganztagsschule und wie definieren sie ihre eigene Rolle?
3. Welche Gelingensbedingungen und Herausforderungen sehen Schulrät*innen in der Implementation von Ganztag in Schulen?

Die Ergebnisse basieren auf leitfadengestützten Interviews mit 12 Schulrät*innen, die zwischen Mai und August 2022 zu ihrer Beratungstätigkeit im Rahmen von Ganztag befragt wurden. Konkret wurde erfragt, welche Rolle das Thema Entwicklung zur Ganztagsschule in ihrer Arbeit spielt, ob es Teil der Beratungsarbeit ist und welchen Beratungsbedarf sie zum Thema Ganztag wahrnehmen. Zudem wurden eigene Tätigkeiten zum Thema Ganztag erfragt und es wurde um eine Einschätzung gebeten, wer die Verantwortung in der Entwicklung zur Ganztagsschule trägt. Gesprächsanlässe waren offen formuliert, sodass eigene Erlebnisse und Überzeugungen geäußert werden konnten. Für die Bearbeitung der Forschungsfragen wurden die Interviews aufgezeichnet, transkribiert und inhaltsanalytisch (Mayring, 2010) in MAXQDA ausgewertet. Dabei wurden zunächst auf dem Leitfaden beruhende deduktive Kategorien formuliert, die induktiv erweitert wurden.

Bei der Betrachtung der Überzeugungen der Schulrät*innen wird deutlich, dass diese überwiegend ausschließlich den gebundenen und verzahnten Ganztag als gelungene Umsetzung betrachten und entkoppelte Hortbetreuung nicht als eine gute Umsetzung von Ganztag ansehen. Ganztag wird als Voraussetzung für die individuelle Förderung einer heterogenen Schülerschaft gesehen, die flächendeckend eingeführt werden sollte. Ein Schulrat aber sah keine Notwendigkeit der Umsetzung von Ganztag an Gymnasien, weil die Unterrichtszeiten aus seiner Sicht lang genug sein.

Die Verantwortung in der Entwicklung zur Ganztagsschule schreiben die Befragten unterschiedlichen Akteur*innen zu und zeichnen – fallübergreifend betrachtet – ein Bild der multiprofessionellen Verantwortung im System Schule, welches Verantwortung auf der Makro-, Meso- und Mikroebene beschreibt und dabei die Rolle der Ministerien, Schulämter, Landesinstitute, Schulleitungsteams, Lehrkräfte, Schüler*innen und Eltern in den Blick nimmt.

Für die Gelingensbedingungen und Herausforderungen zeigt sich, dass die Schulrät*innen insbesondere Herausforderungen in der Unterrichtsversorgung sowie fehlende Ressourcen als Gründe für Schwierigkeiten in der Umsetzung von Ganztag benennen. Als förderlich nehmen sie eine strukturierte inhaltliche Entwicklung mit Fortbildungs- und Unterstützungsangebote von Entwicklungsprogrammen, z. B. von Stiftungen, wahr.

Die Ergebnisse verdeutlichen, welche Überzeugungen Mitarbeitende der Schulaufsicht zum Thema Ganztag haben und wie sie ihre sowie die Rolle unterschiedlicher Akteur*innen in der Entwicklung zu Ganztagsschulen definieren. Außerdem illustrieren sie, vor welchen konkreten praktischen Herausforderungen Schulen in der Entwicklung zur Ganztagsschule stehen und es wird zusammengefasst, welche Unterstützungsangebote in der Bewältigung dieser Herausforderungen helfen könnten.

Wie wirken sich berufliche Gratifikationskrisen auf die Gesundheit von Schulleitungen aus? Untersuchung des Zusammenhangs vom Effort-Reward Imbalance (ERI) Modell und Burnout

Nele Groß[1], Kevin Dadaczynski[2], Marcus Pietsch[1]

[1]Leuphana Universität Lüneburg, Deutschland; [2]Hochschule Fulda, Deutschland

Theoretischer Hintergrund

Die psychische Gesundheit stellt im Sinne der Weltgesundheitsorganisation (2005) eine wesentliche Dimension eines umfassenden Gesundheitsverständnisses dar. Aktuelle Erkenntnisse weisen darauf hin, dass psychische Belastungen und Beanspruchungen im Arbeitskontext eine ernstzunehmende Public Health Herausforderung darstellen. Auf Basis einer Analyse der betrieblichen Fehlzeiten der großen gesetzlichen Krankenkassen kommt die Bundespsychotherapeutenkammer (BPtK, 2018) zu dem Ergebnis, dass sich die Anzahl an Versicherten, die innerhalb eines Jahres aufgrund einer psychischen Erkrankung arbeitsunfähig waren, zwischen den Jahren 2000 und 2017 von 3,3% auf 5,5% erhöht hat. Hasselhorn und Nübling (2004) konnten auf Basis von Daten der BIBB/IAB-Strukturerhebung feststellen, dass das Risiko für das Auftreten arbeitsbedingter psychischer Erschöpfung im Vergleich von 67 Berufsgruppen bei Lehrkräften besonders hoch ausfällt. Auch die Daten der BIBB/BAuA-Erwerbstätigenbefragung 2006 und 2012 weisen auf eine hohe psychische Erschöpfung für Lehrkräfte, die für Angehörige anderer sozialer Berufe jedoch ebenso hoch ausfällt (Cramer et al., 2014). Während für Lehrkräfte mittlerweile zahlreiche Befunde zur psychischen Gesundheit vorliegen, steht die empirische Gesundheitsforschung bei Schulleitungen noch an den Anfängen. Ergebnisse einer aktuellen deutschen Studie zeigen, dass es 45% der befragten Schulleitungen und Schulleitungsmitglieder am Ende eines Arbeitstages schwerfällt, sich zu erholen und ein Drittel fühlt sich oft/immer psychisch erschöpft (Dadaczynski et al., 2021). Zur Analyse der arbeitsbezogenen Bedingungen der psychischen Gesundheit hat sich das Effort-Reward Imbalance (ERI) Modell (Siegrist, 1996a) als geeignet erwiesen. Grundannahme des Modells ist das Prinzip der sozialen Reziprozität. Berufliche Verausgabung (effort) sind Teil eines gesellschaftlich organisierten Austausches, dem berufliche Belohnungen (reward) folgen. Kommt es zu einem Ungleichgewicht zwischen beruflicher Verausgabung und Belohnung, kann diese zu berufliche Gratifikationskrisen mit psychischen Erkrankungen fördernden Stress führen (Siegrist, 1996b). Übersichtsarbeiten konnten zeigen, dass mittels ERI gemessenen Arbeitsbelastungen gute Prädiktoren für psychische Erkrankungen wie beispielsweise Depressionen sind (Siegrist & Dragano, 2008)

Fragestellung

Aufgrund des Mangels an Erkenntnissen zu psychischen Belastungen und Beanspruchungen bei insbesondere Schulleitungen soll der geplante Beitrag die Frage beantworten, inwieweit ein Ungleichgewicht zwischen beruflicher Verausgabung und Belohnung Symptome von Burnout bei Schulleitungen beeinflusst.

Methode

Die analysierten Daten stammen aus der Studie Leadership in German Schools (LineS2020), welche an den Universitäten Tübingen und Lüneburg sowie an der Pädagogischen Hochschule der Fachhochschule Nordwestschweiz durchgeführt wurde. An der für Deutschland repräsentativen Stichprobe nahmen insgesamt 405 zufällig ausgewählte Schulleitungen allgemeinbildender Schulen teil, welche von September bis November 2019 durch die forsa GmbH mithilfe eines Online-Fragebogens befragt wurden. In den Analysen soll unter anderem der Zusammenhang zwischen

dem, ursprünglich als Erklärungsansatz für negative Gesundheitsauswirkungen von Arbeitsbedingungen auf Gesundheit, von Siegrist (1996) entwickelten, Effort-Reward Imbalance (ERI) Modell (Beispielitem: „ Wenn ich an all die erbrachten Leistungen und Anstrengungen denke, halte ich die erfahrene Anerkennung für angemessen.“) und Items aus dem Maslach Burnout Inventory (MBI), wie beispielsweise „Seit ich Schulleiterin/Schulleiter bin, bin ich gleichgültiger gegenüber Menschen geworden.“ (Maslach et al., 1996) Burnout untersucht werden. Unter Einbeziehung von weiteren Variablen soll sukzessive ein multiples Regressionsmodell aufgebaut werden.

Ergebnisse

Erwartet wird ein Zusammenhang zwischen Imbalance beruflicher Verausgabung und Belohnung Symptome von Burnout bei Schulleitungen. Diese Annahme wird zunächst mittels Zusammenhangsanalysen mittels SPSS analysiert und anschließend sukzessive durch multiple Regressionsanalysen ein Regressionsmodell aufgebaut, in dem weitere Hintergrundmerkmale, wie beispielsweise Alter, Geschlecht und Dauer im Schuldienst berücksichtigt werden. Um den Einfluss der fehlenden Werte nicht zu unterschätzen wird das Modell anschließend in MPlus, unter Berücksichtigung der fehlenden Werte durch das Full-Information-Maximum-Likelihood-Verfahren (FIML), analysiert. Sollten sich in den Analysen erwartungskonforme geschlechtsspezifische Unterschiede zeigen, wird nach einer Überprüfung der konfiguralen Messinvarianz ein Gruppenvergleich in MPlus durchgeführt.

Schulleitungshandeln und Kooperationsklima als Prädiktoren von Arbeitszufriedenheit und Mitarbeiterbindung an Schulen

Kris-Stephen Besa, Christoph Kruse, Johanna Gesang, Martin Rothland
Universität Münster, Deutschland

Obgleich die Kooperation an Schulen ein häufig erforschtes Themenfeld pädagogischer Forschung ist, zeigen sich nach wie vor zahlreiche blinde Flecken und divergierende Ergebnisse sowohl mit Blick auf Voraussetzungen für wie auch Auswirkungen von Kooperation. Die Einschätzung, dass die kollegiale Zusammenarbeit von LehrerInnen sowie zwischen Lehrkräften und dem weiteren (pädagogischen) Personal an Schulen bedeutsam ist, wird nicht nur in der Forschung zur Kooperation im LehrerInnenberuf wiederholt betont, sondern auch von den Lehrkräften grundsätzlich geteilt (Fussangel & Gräsel, 2012). Lehrkräfte verfügen vielfach über Erfahrungen hinsichtlich unterschiedlicher kooperativer Praktiken im Beruf und sehen sowohl die innerschulische Zusammenarbeit als auch außerschulische Kooperation als gewinnbringend an (Richter & Pant, 2016). Für die Lehrkräfte selbst werden neben den kollektiven Wirkungen kooperativer Schul- und Unterrichtspraxis die je individuellen Wirkungen wie motivational oder emotional positive Effekte hervorgehoben. So kann den kollegialen Beziehungen – auch als berufliche Entlastungs- und Schutzfaktoren – ein positiver Effekt auf der individuellen Ebene dann zukommen, wenn durch die Kooperation des pädagogischen Personals an Schulen insgesamt eine höhere Arbeitszufriedenheit und Verbundenheit mit der Arbeitsstelle erwächst (vgl. Fussangel & Gräsel, 2012; Massenkeil & Rothland, 2016). Damit Kooperation in Schulen durchgeführt werden kann, sind insbesondere Schulleitungen gefordert, über die Schaffung von Räumen und Zeiten sowie die persönliche Ermutigung zur Kooperation ein angemessenes Kooperationsklima zu schaffen (Kelchtermans, 2006;

Vangrieken, Dochy, Raes & Kyndt, 2015). Welche Zusammenhänge zwischen Kooperation und Arbeitszufriedenheit von Lehrkräften bestehen, ist trotz zahlreicher empirischer Studien aufgrund divergierender Ergebnisse weitgehend unklar.

Der Vortrag wirft einen Blick auf dieses Desiderat durch einen multimethodischen Zugang auf Basis von Daten der Hamburger Schulinspektion. In einer quantitativen Teilstudie mit Daten von 1.052 PädagogInnen an Hamburger Schulen durchgeführte Strukturgleichungsmodellierungen mit akzeptablen Model-Fits (CFI = .90, SRMR = .06, RMSEA = .07) zeigen die Auswirkungen des kooperationsbezogenen Schulleitungshandelns auf Arbeitszufriedenheit und Mitarbeiterbindung, die zum Teil über das schulische Kooperationsklima vermittelt werden. Dabei können für die abhängigen Variablen hohe Varianzaufklärungen von $R2$ = .45 – 68 erzielt werden. Eine qualitative Teilstudie beleuchtet in Ergänzung und Vertiefung Gründe hierfür und identifiziert Satisfiers und Dissatisfiers im Kontext schulischer Kooperation.

A-10: Gendergerechtigkeit

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S05 T00 B83
Chair der Sitzung: Anna K Nishen

Berufswünsche in der 10. Klasse: Vermittelt intrinsische Motivation den Zusammenhang von Geschlecht und Mathenoten mit mathebezogenen Berufswünschen?

Anna K Nishen, Anne Weidinger, Ece Cetin, Ricarda Steinmayr
TU Dortmund, Germany

Mit dem Erreichen eines Abschlusses rückt für Schüler*innen die Entscheidung über ihre Weiterbildung näher und ihrer Berufswünsche werden bedeutsamer. Bestehende Forschung hat gezeigt, dass sowohl motivationale als auch demographische Faktoren eine Rolle bei Berufswünschen von Jugendlichen spielen (Lauermann et al., 2015, 2017). Gerade in Bezug auf Berufe im STEM-Bereich zeigt sich, dass Mädchen bei gleicher Leistung weniger intrinsisch motiviert sind und Berufe mit starkem Mathematikbezug weniger häufig wählen (Lauermann et al., 2015, 2017). Um sowohl die Entstehung von Berufswünschen generell als auch Geschlechterunterschiede in ihnen besser zu verstehen, wird kontextspezifische Forschung benötigt, die potenzielle Mediatoren untersucht (Eccles & Wigfield, 2020). Ein vermittelnde Variable könnte nach dem Erwartungs-Wert-Modell von Eccles et al. (1983) die fachspezifische intrinsische Motivation der Schüler*innen sein, die mit Berufswünschen in Verbindung steht (z.B. Lauermann et al., 2015). Deswegen wurde in der vorliegenden Studie getestet, ob (1) die Berufswünsche von Mädchen im Vergleich zu Jungen ein geringeres Mathematikniveau voraussetzen und ob (2) die intrinsische Motivation für Mathematik diesen Effekt vermittelt.

Die vorliegende Studie nutzt einen bestehenden Datensatz mit n = 1092 Schüler*innen aus 57 zehnten Klassen auf Real-, Haupt- und Gesamtschulen (48.9% weiblich, 82% Deutsch als Erstsprache, M_{Alter} = 16.09, SD = 0.68). Es wurde das Geschlecht und die intrinsische Motivation in Mathematik sowie in einem offenen Format der Berufswunsch erfasst, der verschiedenen Berufsgruppen des O*NET zugeordnet wurde (Version 25.2; https://www.onetonline.org). Die Kodierung nach dem O*Net erlaubt die Bestimmung des ebenfalls in der Datenbank hinterlegten Niveaus an Mathematikkompetenzen, das für die Ausübung der Berufe in den jeweiligen Berufsgruppen benötigt wird. Dieses benötigte Mathematikniveau diente als abhängige Variable. Mithilfe einer Pfadanalyse wurden der Effekt des Geschlechts auf das Mathematikniveau im Wunschberuf in der 10. Klasse unter Kontrolle der letzten Zeugnismathematiknote, Mathematikkompetenzen (KRW, Schmidt et al., 2013; Kurzform des TIMSS-Tests, Lauermann et al., 2020) und fluider Intelligenz (CFT 20-R, Weiß, 2006), die zu einem früheren Testzeitpunkt erhoben wurden, gerechnet. Die genestete Datenstruktur wurde anhand von z-Standardisierung der Kontrollvariablen innerhalb von Klassen berücksichtigt. Darüber hinaus wurde in einer Mediatoranalyse mittels des SPSS-Makros PROCESS (Hayes, 2022) getestet, ob die intrinsische Motivation diesen Effekt vermittelte.

Unseren Hypothesen entsprechend zeigten die Ergebnisse, dass das Geschlecht damit zusammenhängt, wie hoch das Mathematikniveau des gewünschten Berufes von Zehntklässler*innen ist (b = −0.34, SE = 0.07, p < .001, β = −0.18). Die Berufswünsche von Mädchen wiesen ein geringeres benötigtes Mathematikniveau auf. Darüber hinaus zeigten Mädchen eine geringere intrinsische Motivation für Mathe (b = −0.21, SE = .06, p = .001, β = −0.11). Diese hing wiederum

positiv mit der Höhe der benötigten Mathekompetenz im Wunschberuf zusammen (b = 0.10, SE = 0.04, p = .011, β = 0.10). Der indirekte Effekt auf das Mathematikniveau des Berufswunsches vermittelt durch die intrinsische Motivation war ohne Kontrolle der Kovariaten signifikant (−0.07, SE = 0.02; 95%-KI [−0.11, −0.03]. Das war auch nach Kontrolle der oben genannten Kovariaten noch der Fall, auch wenn der Effekt sich verringerte (b = −0.03, SE = 0.02, 95%-KI [−0.07, −0.01]).

Die Ergebnisse zu Geschlechterunterschieden sind im Einklang mit bestehender Forschung im STEM-Bereich (Lauermann et al., 2015; Gaspard et al., 2020). Mädchen fühlen sich trotz geringer Geschlechtsunterschiede in den mathematischen Kompetenzen im Schnitt weniger zu Berufen hingezogen, die ein höheres mathematisches Niveau erfordern, als Jungen. Die Ergebnisse deuten darauf hin, dass die intrinsische Motivation und Berufswünsche von Jugendlichen weiterhin von Geschlechterrollen geprägt sind. Die intrinsische Motivation mediierte den Effekt von Geschlecht auf das Mathematikniveau der Berufswünsche jedoch nur teilweise, sodass weitere Mediatoren in zukünftigen Analysen in Betracht gezogen werden sollten. Dies sollte in zukünftigen Studien unter Berücksichtigung weiterer Variablen wie Geschlechtsrollenselbstkonzept untersucht werden. Limitationen wie die querschnittliche Natur der Daten werden diskutiert.

Bullying in der Schule: Sind die Hilfemaßnahmen angehender Lehrkräfte abhängig vom Geschlecht der Schüler*innen?

Sabine Glock[1], Hannah Kleen[2]
[1]Bergische Universität Wuppertal, Deutschland; [2]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Theoretischer Hintergrund

Aggression und Gewalt unter Schüler*innen ist immer noch ein Thema in der Schule und vor allem angehende Lehrkräfte fühlen sich überfordert und unzureichend vorbereitet, angemessen auf Bullying zu reagieren (Bilz et al., 2019). Erfahrene Lehrkräfte scheinen eher zu Maßnahmen zu neigen, die ihre Autorität betonen, wie Zurechtweisungen oder auch Sanktionen (Burger et al., 2015). Insgesamt scheint die Anzahl der Bullyingereignisse abzunehmen, aber die Häufigkeit, mit der Mädchen oder Jungen in das Geschehen verwickelt sind hat sich verändert (Niproschke et al., 2019). Vor dem Hintergrund der Geschlechterstereotype als Wissen(Smith, 1998) darüber, wie Jungen und Mädchen sich verhalten sollten und es auch tun (Prentice & Carranza, 2002), müssten Unterschiede für das Handling von Bullyingsituationen zu erwarten sein, auch wenn die Auftretenshäufigkeit von Bullying sich kaum mehr zwischen Jungen und Mädchen unterscheidet (Niproschke et al., 2019). So werden Mädchen stereotyp mit den Eigenschaften brav, fleißig und angepasst in Verbindung gebracht (Glock & Kleen, 2017), während Jungen als wenig angepasst und wenig selbstdizipliniert wahrgenommen werden (Hannover & Kessels, 2011).

Fragestellung

Die Frage, die sich vor diesem Hintergrund stellt, ist, welche Interventionen angehende Lehrkräfte als Reaktion auf eine Bullyingsituation verwenden und es hier Unterschiede in Bezug auf das Geschlecht der Schüler*innen gibt. Diese Fragestellungen wurden in einer experimentellen Vignettenstudie getestet.

Methode

An der Studie nahmen 334 Lehramtsstudierende aller Studiengänge teil. Es wurden zwei Vignetten zusammengestellt, eine in der sowohl Täterin als auch Opfer weiblich waren, und eine in der beide Hauptfiguren männlich waren. Folgendes wurde in den Vignetten beschrieben: „Als ein*e bestimmte Schüler*in sich einer Gruppe nähert, ruft ein*e Schüler*in laut: ‚Hau ab, niemand will dich dabeihaben'". Als Interventionsmaßnahmen wurde die Liste von Burger et al. (2025) benutzt, die in die fünf Maßnahmen autoritäts-basierte Interventionen, Ignorieren, Arbeit mit dem Opfer, Arbeit mit dem Täter und das Hinzuziehen anderer Erwachsener gegliedert sind. Die Lehramtsstudierenden gaben auf einer Skala von 1 (auf keinen Fall) bis 5 (auf jeden Fall) an, wie wahrscheinlich sie die Maßnahmen anwenden würden.

Ergebnisse und ihre Bedeutung

Eine ANOVA mit dem Zwischengruppenfaktor Schüler*innengeschlecht und dem messwiederholten Faktor Interventionsmaßnahmen ergab einen signifikanten Haupteffekt der Interventionsmaßnahmen $F(4, 1316) = 323.90$, $p < .001$, $\eta_p^2 = 0.50$), der zeigte, dass generell die wahrscheinlichsten Maßnahmen autoritäts-basiert ($M = 4.18$, $SD = 0.51$) und die Arbeit mit dem*r Täter*in waren ($M = 3.08$, $SD = 0.69$). Die Interaktion zwischen Schüler*innengeschlecht und Interventionsmaßnahmen wurde ebenfalls signifikant, $F(4, 1316) = 3.22$, $p < .012$, $\eta_p^2 = 0.01$. Das autoritäre Auftreten ($M_{Jungen} = 4.11$, $SD_{Jungen} = 0.68$; $M_{Mädchen} = 3.92$, $SD_{Mädchen} = 0.70$) und auch das Hinzuziehen anderer Erwachsener wurde wahrscheinlicher bei Jungen ($M = 3.04$, $SD = 0.91$) angegeben als bei Mädchen ($M = 2.82$, $SD = 0.82$). Die anderen Maßnahmen unterschieden sich nicht, auch der Haupteffekt Schülergeschlecht wurde nicht signifikant.

Die Ergebnisse zeigen, dass angehende genau wie erfahrene Lehrkräfte auf Bullying mit autoritären Maßnahmen reagieren, aber auch mit der*m Täter*in arbeiten würden, ohne direkt zu strafen. Dabei scheinen sie die Opfer stärker außer Acht zu lassen, obwohl diese auch auf Unterstützung der Lehrkräfte angewiesen sind. Für Jungen werden wahrscheinlicher autoritäre Maßnahmen ergriffen und Hilfe bei anderen Erwachsenen gesucht als bei Mädchen. Dies kann daran liegen, dass Lehrkräfte eher denken, dass das Ausgrenzungsverhalten bei Mädchen etwas ist, das Mädchen selbstständiger lösen können als Jungen, da sie als Vertreter der weiblichen Personen mit sozialer Kompetenz in Verbindung gebracht werden (Steffens & Ebert, 2016). Die Maßnahmen für die Jungen scheinen Hilflosigkeit zu signalisieren. Unabhängig von den Hintergründen der Ergebnisse, impliziert die Studie, dass Unterschiede Abhängigkeit des Geschlechts zu einer Ungleichbehandlung in der Schule führen können.

Aber das gendert doch nichts – Welchen Einfluss haben gender-gerechte Sprachalternativen auf den kognitiven Einbezug von Frauen?

Christin Lotz, Anne Deiglmayr
Universität Leipzig, Germany

Theoretischer Hintergrund: Die Verwendung gender-gerechter Sprache wird im deutschen Sprachraum intensiv diskutiert. Der dabei häufig vorherrschenden populär-wissenschaftlichen Debatte um die Wirksamkeit gender-gerechter Sprache auf den kognitiven Einbezug von Frauen und Personen diversen Geschlechts steht allerdings nur eine geringe Anzahl an methodisch belastbaren empirischen Studien gegenüber, die jedoch ermutigende Ergebnisse zeigten (z.B. Braun et al., 2005;

Gabriel et al., 2018; Sczesny et al., 2016; Horvath et al., 2016). Unabhängig von der Frage, *ob* gegendert werden sollte, stellt sich jedoch insbesondere auch die Frage, *wie* gegendert werden sollte. Diese Entscheidung basierte in der Vergangenheit häufig auf nicht-empirischen Überlegungen. Während beispielsweise die Nennung beider Geschlechter (z.B. Lehrerinnen und Lehrer) impliziert, dass Geschlechtsidentität binär ist, stellen Gender-Zeichen wie der Doppelpunkt (z.B. Lehrer:innen) einen Platzhalter für Personen jenseits des binär männlichen oder weiblichen Geschlechts dar. Des Weiteren werden unter dem Aspekt der Barrierefreiheit pragmatische Gründe wie z.B. der möglichst störungsfreie Vorlesefluss gegenderter Texte von Screenreadern angeführt. Empirische Studien, die die Effekte der unterschiedlichen gender-gerechten Sprachalternativen im Vergleich zum generischen Maskulinum und untereinander auf den kognitiven Einbezug von Frauen untersuchten, sind bislang kaum vorhanden oder bezogen nur sehr wenige gender-gerechte Sprachalternativen ein (Braun et al., 2005; Koch, 2021).

Fragestellung: Die folgende Studie untersuchte daher umfassender den Einfluss von acht gängigen gender-gerechten Sprachalternativen (a) im Vergleich zum generischen Maskulinum und (b) im Vergleich zueinander in ihrem Effekt auf den kognitiven Einbezug von Frauen.

Methode: Dafür wurden N = 817 Studierenden (n = 356 männlich; n = 461 weiblich) per Online-Fragebogen sechs Oberkategorien (Politiker, Sänger, Sportler…) in unterschiedlichen Sprachformen präsentiert, zu denen sie drei berühmte Personen nennen sollten (vgl. Braun et al., 2005). Im between-subjekt design wurden die Studierenden zufällig einer der neun Sprachformen-Bedingungen zugewiesen (Generisches Maskulinum: Politiker; Beidnennung, männlich zuerst: Politiker und Politikerinnen; Beidnennung, weiblich zuerst: Politikerinnen und Politiker; Gender-Sternchen: Politiker*innen; Gender-Doppelpunkt: Politiker:innen; Binnen-I: PolitikerInnen; Unterstrich: Politiker_innen; verkürzte Beidnennung mit Klammern: Politiker(innen); verkürzte Beidnennung mit Schrägstrich: Politiker/-innen). Basierend auf einer männlich-weiblich-Kodierung berechneten wir als abhängige Variable die relative Häufigkeit weiblicher Personen-Nennungen an der Gesamtzahl genannter Personen (Wertebereich von 0 [keine weibliche Person genannt] bis 1 [nur weibliche Personen genannt]). Anschließend berechneten wir eine 9 (Sprachform) × 2 (Geschlecht) ANOVA mit anschließenden LSD Post-hoc Tests.

Ergebnis: Die ANOVA ergab einen signifikanten Haupteffekt für die Sprachform, einen signifikanten Haupteffekt für das Geschlecht der Teilnehmenden, aber keinen signifikanten Interaktionseffekt. Der Haupteffekt Sprachform implizierte, dass sich die Verwendung der gender-gerechten Alternativen zum generischen Maskulinum positiv auf den kognitiven Einbezug von Frauen auswirkte. Post-hoc Tests ergaben, dass alle alternativen Sprachformen außer der verkürzten Beidnennung mit Klammern zu einer häufigeren Nennung von weiblichen Personen führten als die Sprachform des generischen Maskulinums ($0.24 \leq d \leq 0.45$). Ein Vergleich der gender-gerechten Sprachformen untereinander ergab, dass sich die Häufigkeit der Nennung weiblicher Personen nicht signifikant über die Sprachformen hinweg unterschied (einzige Ausnahme: Unterstrich > verkürzten Beidnennung mit Klammern, d = 0.33). Der Haupteffekt Geschlecht zeigte an, dass Frauen mehr weibliche Personen nannten als Männer (d = 1.42).

Diskussion: Die Verwendung gender-gerechter Sprachalternativen zum generischen Maskulinum hat einen positiven Effekt auf den kognitiven Einbezug von Frauen. Aus empirischer Sicht lässt sich dabei keine explizite Empfehlung für die Verwendung einer bestimmten alternativen Sprachform zum generischen Maskulinum aussprechen. Viel mehr scheint es wichtig zu sein, dass gender-gerechte Sprachalternativen überhaupt genutzt werden. Es erscheint daher sinnvoll, dass die vorliegenden Ergebnisse mit zukünftigen Studien zur Akzeptanz, Verständlichkeit und Praktikabilität der gender-gerechten Sprachalternativen komplementiert werden, um eindeutige praktische Empfehlungen für oder ggf. gegen eine gender-gerechte Sprachalternative aussprechen zu können.

Komposition und Kultur der Schüler:innen und Lehrkräfte und die Entwicklung von Geschlechterrollen im Jugendalter

Ricarda Ullrich[1], Mieke Van Houtte[2], Michael Becker[3,1]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; [2]Universität Gent, Belgien; [3]Institut für Schulentwicklungsforschung (IFS), TU Dortmund, Deutschland

Geschlechterrollen können als normative Erwartungen über die geschlechtliche Macht- und Arbeitsverteilung verstanden werden. Traditionell wird der männlichen Rolle die Aufgabe des „Brotverdienens“ zugewiesen, während weibliche Personen für häusliche und reproduktive Tätigkeiten verantwortlich sind (Eagly & Wood, 2012). Jugendliche sind diesen Geschlechterrollen ausgesetzt und entwickeln hierbei ihre eigenen geschlechtsbezogenen Einstellungen. Diese Einstellungen verändern sich während der Adoleszenz (Halimi et al., 2021; Ullrich et al., 2022), da Jugendliche neuen sozialen Kontexten ausgesetzt sind, die sich auf ihre Einstellungen auswirken können (Katz & Ksansnak, 1994), und obwohl der Schulkontext als wichtiger Sozialisationskontext erachtet wird (Bronfenbrenner, 1977), wird dieser in nur wenigen Studien eigens berücksichtigt. Die Lehrkräfte als wichtige Akteure im Schulkontext werden zudem häufig übersehen (Van Houtte, 2011).

In den wenigen Arbeiten, die bislang hierzu vorgelegt wurden, zeigte sich, dass die Varianz in Geschlechterrolleneinstellungen größer zwischen Schulen als innerhalb von Schulen ist, was eine geschlechtsrollenbezogene Kultur innerhalb von Schulen nahelegt, die wiederum mit Kompositionsmerkmalen der Schule variiert (Van Houtte, 2021; Vantieghem, 2015). Da weibliche Personen egalitärere Geschlechterrolleneinstellungen zeigen, sollte die Geschlechtsrollenkultur in Schulen mit mehr Schülerinnen und Lehrerinnen offener und egalitärer sein (Huyge et al., 2015; Vantieghem, 2015). Darüber hinaus ist ein privilegierterer familiärer sozioökonomischer Status mit weniger traditionellen Geschlechterrollen bei Jugendlichen assoziiert, jedoch nicht mit der Entwicklung von jugendlichen Geschlechterrolleneinstellungen (Halimi et al., 2021; Ullrich et al., 2022). Anzunehmen ist daher, dass die sozioökonomische Komposition der Schüler:innen wichtiger für die Entwicklung von Geschlechterrollen sein könnte.

Lehrkräfte haben ebenfalls internalisierte Geschlechterrolleneinstellungen, welche durch geschlechtsbezogene Erwartungen an die Schüler:innen weitergetragen werden können (Wolter et al., 2015). Durch ihre Vorbildfunktion und die Vermittlung von geschlechtsspezifischen Erwartungen können Lehrkräfte bewusst oder unbewusst die gesellschaftliche Geschlechterverteilung reproduzieren. Dabei zeigte Van Houtte (2011), dass bestimmte Erwartungskulturen innerhalb von Lehrkräften einer Schule entstehen, die darauf hindeuten, dass auch eine Lehrerkultur bezüglich

Geschlechterrollen entstehen könnte. Diese Lehrerkultur könnte wiederum Einfluss auf die Geschlechterrollen von Schülern und Schülerinnen haben.

Entsprechend wird in dem Beitrag der Frage nachgegangen, inwieweit die Geschlechterrollenkultur, die Geschlechterzusammensetzung und die SES-Zusammensetzung der Schüler:innen sowie die Geschlechterrollenkultur und die Geschlechterzusammensetzung der Lehrkräfte die Entwicklung der Geschlechterrolleneinstellungen von Schüler und Schülerinnen im Jugendalter erklären können. Drei Wellen eines flämischen Datensatzes (N=6311, Schulen=59) werden verwendet, um die Entwicklung von Geschlechterrolleneinstellungen kurz nach dem Übergang in die Sekundarstufe zwischen der siebten und achten Klasse zu erklären (2012-2014, erste und zweite Klasse der Sekundarstufe). Die Analysen werden als Mehrgruppen Baseline Change Modelle mit der *type =complex* Option in Mplus 8.4 durchgeführt.

Die Ergebnisse zeigen, dass insbesondere Schüler empfänglich sind für die kompositionellen und kulturellen Merkmale der Schüler:innen und Lehrkräfte. Eine traditionellere Geschlechterrollenkultur der Schüler:innen geht bei Jungen mit traditionelleren Einstellungen zur siebten Klasse einher. Allerdings entwickeln sich die Einstellungen dieser Schüler, die bereits zu Beginn der Sekundarstufe traditionellere Geschlechterrolleneinstellungen zum Ausdruck bringen, nicht so stark in eine noch traditionellere Richtung. In Kontexten mit einer privilegierteren sozioökonomischen Zusammensetzung entwickeln Jungen weniger traditionelle Einstellungen, während eine traditionellere Geschlechterrollenkultur der Lehrkräfte die Entwicklung traditionellerer Geschlechterrolleneinstellungen bei Schülern fördert. Für die Entwicklung der Mädchen wurden keine signifikanten Effekte der kulturellen und kompositorischen Faktoren der Schüler:innen und Lehrkräfte festgestellt. Werden Jungen und Mädchen verglichen, dann zeigen die Ergebnisse, dass Jungen sensitiver sind für die Kultur und Zusammensetzung von Lehrkräften und Schüler:innen, wenn es um die Entwicklung von Geschlechterrolleneinstellungen geht, was die Bedeutung der Untersuchung geschlechtsspezifischer Kulturen im schulischen Kontext unterstreicht.

Symposien

A-11: Die Entwicklung von Schüler:innen im Verlauf der COVID-19-Pandemie – Ein Stand 3 Jahre nach Pandemiebeginn

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* S07 S00 D07

Die Entwicklung von Schüler:innen im Verlauf der COVID-19-Pandemie – Ein Stand 3 Jahre nach Pandemiebeginn

Chair(s): **Antje Ehlert** (Universität Potsdam, Deutschland)
*Diskutant*in(nen):* **Kai Maaz** (Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF), Frankfurt)

Die COVID-19-Pandemie zwang viele Länder zu abrupten Schulschließungen und meist zur Aufnahme des Distanzunterrichts. Diese sehr herausfordernde Zeit führte allerdings zu Lernrückständen und zur Vergrößerung bestehender Bildungsungleichheiten. So berichten Hammerstein et al. (2021) für die Schulschließung 2020 Auswirkungen, die vergleichbar mit dem Lernverlust während der Sommerferien sind (s.a. Kuhfeld et al., 2021). Insbesondere jüngere Kinder (Tomasik et al., 2021) und Schüler:innen aus Familien mit niedrigem sozioökonomischen Status (Engzell et al., 2021; Maldonado & De Witte, 2020) sind negativ von den Schulschließungen und dem Distanzunterricht betroffen. International und national kommen in Abhängigkeit von der elterlichen Bildung und dem Einkommen somit weitere Lernverluste zu den bereits bestehenden Ungleichheiten in Lesen, Rechtschreiben und Mathematik hinzu (z.B. Haerlermans et al., 2022; Stanat et al., 2022; Ludewig et al., 2022). Neben den Lernverlusten erhöhte sich auch die Streuung der Kompetenzwerte (Stanat et al., 2022), und soziale Disparitäten verstärkten sich in allen Fächern und Kompetenzbereichen, insbesondere für Heranwachsende mit Migrationshintergrund und ungünstigen häuslichen Lernbedingungen (Goudeau et al., 2021). Ergebnisse zeigen, dass der Distanzunterricht und dessen Gelingen abhängig von der elterlichen Unterstützung, den familiären sowie individuellen Ressourcen und den Lehrpersonenkompetenzen ist (Engzell et al., 2021; Helm et al., 2021; Schult et al., 2022).

Das Symposium hat zum Ziel, zum einen Schüler:innenentwicklungen über die Zeit der Pandemie bis jetzt zu beleuchten, beeinflussenden Faktoren während der Schulschließungen zu analysieren und zum anderen einen Ausblick auf den Umgang mit diesen Ergebnissen zu geben.

Der erste Beitrag untersucht, wie sich lernbezogene Schüler:inneneinschätzungen infolge des Distanzunterrichts verändern und Risikofaktoren diese beeinflussen sowie den Zusammenhang zwischen wahrgenommener Lehrpersonenunterstützung und lernbezogenen Schüler:inneneinschätzungen.

Der zweite Beitrag prüft anhand verpflichtender Lernstandserhebungen die Schüler:innenkompetenzen im zweiten und dritten Pandemiejahr, wobei Lernrückstände und –zuwächse quantifiziert und mit sozialen Kontextvariablen in Bezug gesetzt werden.

Der dritte Beitrag analysiert, inwiefern individuelle, familiäre und lehrkraftseitige Aspekte mit der Schüler:innenlernmotivation während der Schulschließung assoziiert sind, wobei die Bedeutung individueller Lernvoraussetzung unter Berücksichtigung von Lehrkraftmerkmalen zur Vorhersage motivationaler Zielkriterien im besonderen Fokus stehen.

Die aktuelle Ergebnislage gibt Hinweise auf Schüler:innengruppen mit besonders hohem Risiko für anhaltende Lernrückstände. Eine steigende Anzahl an Risikomerkmalen scheint sich teilweise

ungünstig auf die Entwicklung lernbezogener Schüler:innenmerkmale auszuwirken. Allerdings wird auch gezeigt, dass Lernrückstände aufholbar sind, wenn leistungsschwache Schüler:innengruppen Unterstützungsangebote im Unterricht erhalten. Ebenso zeigt sich, dass Lehrpersonenunterstützung sowie die Zusammenarbeit mit Eltern eine wichtige Rolle für das Gelingen des Distanzunterrichts darstellt. So geben die Ergebnisse Hinweise auf Unterstützungsmaßnahmen bei besonders gefährdeten Schüler:innen sowie protektive Faktoren für den Unterricht.

Beiträge des Symposiums

Schüler:innen aus Risikogruppen im Distanzunterricht: Veränderungen lernbezogener Merkmale

Anne Hartmann[1], Antje Ehlert[1], Jenny Lenkeit[1], Michel Knigge[2], Nadine Spörer[1]
[1]Universität Potsdam, [2]Humboldt Universität zu Berlin

Theoretischer Hintergrund:

Das Aussetzen des Präsenzunterrichts im Frühjahr 2020 infolge der Covid-19-Pandemie führte zu grundlegenden Veränderungen der Lernsituation vieler Schüler:innen über viele Monate hinweg (Reintjes et al., 2021). Es wird davon ausgegangen, dass die Herausforderungen, die mit dem Distanzunterricht verbunden sind, die Lernprozesse für Schüler:innen erschweren (Engzell et al., 2021; Nusser, 2021). Hohe Lehrkompetenzen sowie differenzierende Maßnahmen der Lehrkraft zur Unterstützung der Schüler:innen werden als wichtige Voraussetzungen für eine qualitätsvolle Gestaltung des Distanzunterrichts betrachtet (Helm et al., 2021). Zudem zeigen Studien, dass das Gelingen des Distanzunterrichts in hohem Maße von vorhandenen individuellen und familiären Ressourcen der Schüler:innen abhängt (Engzell et al., 2021; Nusser, 2021; Schult et al., 2022). So wird angenommen, dass Schüler:innen mit sonderpädagogischem Förderbedarf (SPF) und Schüler:innen aus sozial benachteiligten Elternhäusern von den leistungsbezogenen Folgen des Distanzunterrichts besonders negativ betroffen sind (Goldan et al., 2020; Engzell et al., 2021; Schult et al., 2022). Offen bleibt jedoch, inwiefern sich lernbezogene Einschätzungen von Schüler:innen, wie z.B. das Selbstkonzept, veränderten und ob diese Veränderungen unterschiedlich für verschiedene Schüler:innengruppen ausfallen.

Fragestellungen:

Vor diesem Hintergrund ergeben sich folgende Fragestellungen:

1. Wie haben sich lernbezogene Einschätzungen von Schüler:innen infolge des Distanzunterrichts verändert?

2. Weisen Schüler:innen aus Risikogruppen ungünstigere Veränderungen in Bezug auf lernbezogene Einschätzungen auf als ihre Peers, die nicht zu Risikogruppen gehören?

3. In welchem Zusammenhang stehen die wahrgenommene Unterstützung der Lehrkraft im Distanzunterricht mit lernbezogenen Einschätzungen der Schüler:innen?

Methode:

Die Daten stammen aus einer Längsschnittstudie zur inklusiven Bildung im Bundesland Brandenburg (t1: Juni 2019) und einer Anschluss-Studie zu Veränderungen infolge des Distanzunterrichts (t2:

November 2021). Es liegen Informationen von 638 Schüler:innen aus 21 Klassen der 5. und 14 Klassen der 10. Jahrgangsstufe vor.

Lernbezogene Merkmale sind als fachspezifische Selbstkonzepte und Interessen für Mathematik und Lesen operationalisiert (t1: α = .76-.90; t2: α = .76-.92). Zu t2 wurde zusätzlich die wahrgenommene Unterstützung der Lehrkraft im Distanzunterricht erfragt (α = .74). Darüber hinaus lagen von den Schüler:innen Informationen zum SPF, sozialen Hintergrund, zur gesprochenen Sprache Zuhause sowie zum zusätzlichen Unterstützungsbedarf im Bereich Lernen, Sprache sowie Arbeits- und Sozialverhalten vor. Aus diesen 6 Variablen wurde ein additives Risikomaß berechnet, welches von 0 bis 5 Risikomerkmalen reichte. Mithilfe von Growth-Curve-Modellen wurden Veränderungen in den lernbezogenen Einschätzungen sowie deren Zusammenhänge mit dem Risikomaß und der wahrgenommenen Unterstützung der Lehrkraft geprüft.

Ergebnisse und ihre Bedeutung:

Erste Analysen zeigten vorrangig Abnahmen lernbezogener Einschätzungen der Schüler:innen nach der Phase des Distanzunterrichts (z.B. Interesse für Mathematik: Jahrgang 5: $b = -0.63$, $p < .05$; Jahrgang 10: $b = -0.10$, $p < .05$). In der 5. Jahrgangsstufe zeigte sich, dass mit steigender Anzahl an Risikomerkmalen das fachspezifische Selbstkonzept sowie Interesse für Mathematik und Lesen stärker abnahm. Bei Schüler:innen der 10. Jahrgangsstufe fiel das Selbstkonzept Lesen mit steigender Anzahl an Risikomerkmalen geringer aus ($b = -0.20$, $p < .05$), jedoch stieg das Interesse für Mathematik mit zunehmender Anzahl an Risikomerkmalen ($b = 0.19$, $p < .05$). Zusätzlich zeigte sich in dieser Jahrgangsstufe ein signifikanter Moderationseffekt: Von t1 zu t2 nahm das Selbstkonzept Lesen bei steigender Anzahl an Risikomerkmalen weniger stark ab ($b = 0.18$, $p < .05$). In der 5. Jahrgangsstufe ließ sich ein signifikanter Moderationseffekt der wahrgenommenen Unterstützung im Distanzunterricht über die Zeit hinweg beobachten: Fühlten sich Schüler:innen durch ihre Lehrkraft unterstützt, nahm das Interesse für Mathematik von t1 zu t2 weniger stark ab ($b = 0.34$, $p < .05$).

Diese ersten Ergebnisse deuten darauf hin, dass sich infolge des Distanzunterrichts Disparitäten unter Schüler:innen verstärken haben. Sie legen weiterhin nahe, dass die Unterstützung der Lehrkraft eine wichtige Rolle für das Gelingen des Distanzunterrichts darstellt.

Lernrückstände und -zuwächse im zweiten und dritten Pandemiejahr: Ausmaß und soziale Risikofaktoren

Johannes Schult[1], Nicole Mahler[1], Benjamin Fauth[2], Marlit A. Lindner[3]
[1]Institut für Bildungsanalysen Baden-Württemberg (IBBW), Stuttgart, [2]Institut für Bildungsanalysen Baden-Württemberg (IBBW), Stuttgart, Hector-Institut für Empirische Bildungsforschung, Universität Tübingen, [3]Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Kiel

Während den weltweiten Schulschließungen aufgrund der COVID-19-Pandemie verkürzte sich die Lernzeit der Schülerinnen und Schüler deutlich. Distanzunterricht konnte nicht verhindern, dass die Lernleistungen in Lesen und Mathematik im Vergleich zu Vor-Pandemie-Jahren geringer ausfielen. Die meta-analytische Schätzung des Lernrückstands über Fächer, Klassenstufen und Länder hinweg beträgt d = -0.14 (Betthäuser et al., 2022).

Die Lernrückstände waren besonders hoch ausgeprägt bei Schülerinnen und Schülern mit niedrigem sozio-ökonomischen Status und wenig sozio-kulturellem Kapital (Haelermans et al., 2022). Auch in

Deutschland gab es bei Schulen mit einer hoch belasteten Schülerschaft im Schnitt die größten Lernrückstände (Lücken, 2022). Nach der weitgehenden Rückkehr zum Präsenzunterricht konnten Lernlücken teilweise geschlossen werden (Haelermans et al., 2021). Tendenziell gelang dies am ehesten leistungsstarken Schülerinnen und Schülern (Kuhfeld et al., 2022).

In der vorliegenden Arbeit werden anhand von verpflichtenden Lernstandserhebungen die Kompetenzen von Schülerinnen und Schülern im zweiten und dritten Pandemie-Jahr untersucht. Dabei sollen Lernrückstände und -zuwächse quantifiziert und in Bezug zu sozialen Kontext-Variablen gesetzt werden.

Beim „Lernstand 5“ werden ausgewählte Basiskompetenzen der Bildungsstandards in Deutsch und Mathematik zu Beginn der 5. Klasse erhoben. Die Kompetenztests finden jährlich an allen allgemein bildenden öffentlichen Schulen mit Sekundarstufe I in Baden-Württemberg statt. Aufgrund der verpflichtenden Teilnahme werden die Kohorten jeweils nahezu vollständig getestet (jeweils n > 80,000). Für die Analyse von Lernrückständen und etwaigem Aufholen werden die Pandemiejahre 2020 bis 2022 betrachtet und verglichen mit den Leistungen von 2017 bis 2019. Neben den Mittelwerten werden auch die 5%-, 25%-, 50%-, 75%- und 95%-Perzentile untersucht. Mittels linearer Regressionen wird untersucht, wie weit die Lernrückstände bzw. -zuwächse durch sozio-kulturelles Kapital, Migrationshintergrund sowie die Dauer der Schulschließungen erklärt werden können (Schult et al., 2022).

Für 2021 zeigt sich, dass die ursprünglichen Lernrückstande in Deutsch (Lesen) und Mathematik (Zahlverständnis) teilweise aufgeholt werden konnten. Im Mathematik-Bereich Operationsverständnis nahmen die Lernrückstände hingegen weiter zu, wenn auch nicht mehr so stark wie 2020. Für das leistungsschwächste 5%-Perzentil blieben die Lernrückstände in der Regel am größten. Schulschließungsdauer und niedriges sozio-kulturelles Kapital hingen positiv mit dem Ausmaß der Lernrückstände zusammen. Für verschiedene Ausprägungen des Faktors Migrationshintergrund zeigte sich darüber hinaus jedoch kein bedeutsamer Zusammenhang. Die Datenerhebung von 2022 war zum Zeitpunkt der Einreichung des Beitrags für die GEBF noch nicht abgeschlossen. Die Landesergebnisse 2022 werden aber auf der Konferenz vorgestellt.

Die Befunde geben deutliche Hinweise auf Schülergruppen mit einem besonders hohen Risiko für anhaltende Lernrückstände. Sie helfen, Unterstützungsangebote gezielt zu steuern. Da einzelne Erhebungen, wenn sie auf wenige Punktschätzer reduziert werden, im scheinbaren Widerspruch zu (einzelnen Punktschätzern aus) anderen Studien stehen, werden die Ergebnisse zur weiteren Einordnung vor dem Hintergrund des Bildungstrends 2021 (Stanat et al., 2022) sowie weiterer Lernstandserhebungen diskutiert. Bei jedem spezifischen Studiendesign gibt es bestimmte Besonderheiten und Limitationen. Im meta-analytischen Gesamtbild der Lernrückstände können scheinbare Widersprüche allerdings aufgelöst werden, so dass von einer klaren Evidenz gesprochen werden kann. Die baden-württembergischen Vollerhebungen unterstreichen die Möglichkeit, Lernrückstände aufzuholen und leistungsschwachen Schülergruppen auch im Unterricht Unterstützungsangebote anzubieten, die direkt an die Kompetenzdiagnose anschließen.

Motivation von Schülerinnen und Schülern beim Lernen in Distanz: Das Zusammenspiel individueller, familiärer und lehrkraftseitiger Einflussgrößen

Richard Göllner, Ann-Kathrin Jaekel
Universität Tübingen, Hector-Institut für Empirische Bildungsforschung

Theoretischer Hintergrund

Der schulische Unterricht und dessen Durchführung waren während der Schulschließungen zur Eindämmung der COVID 19-Pandemie in erster Linie dadurch geprägt, dass Unterricht nicht gemeinsam im Klassenzimmer stattfand, sondern zu Hause erfolgen musste. Dabei zeigten sich nicht nur zwischen Lehrkräften erhebliche Unterschiede darin, wie Unterricht umgesetzt wurde, beispielsweise im Hinblick auf Unterrichtsmethoden oder die Bereitstellung von Aufgaben (Jaekel, Scheiter & Göllner, 2021; Vodafone Stiftung, 2020). Auch erforderte der Distanzunterricht ein höheres Maß an Austausch mit den Eltern, um den Unterricht zu organisieren und verlangte von den Lehrkräften spezifische Fertigkeiten im Umgang mit digitalen Medien (z.B. Reimers & Schleicher, 2020). Gleichzeitig ist davon auszugehen, dass die individuellen Lernvoraussetzungen in Form von Lernhaltungen der Schülerinnen und Schüler und der elterlichen Unterstützung den Lernerfolg erheblich mitbestimmten (z.B. Weber, Helm & Kemethofer, 2021). In diesem Zusammenhang zeigten unterschiedliche Studien der vergangenen drei Jahre negative Effekte der Schulschließungen insbesondere für jüngere Schülerinnen und Schüler oder Schülerinnen und Schüler mit einem niedrigeren sozioökonomischen Status (Goudeau et al., 2021; Hammerstein et al., 2021). Vor diesem Hintergrund ging die vorliegende Studie der Frage nach, inwieweit individuelle, familiäre und lehrkraftseitige Einflussgrößen mit der Lernmotivation von Schülerinnen und Schülern während der Schulschließung assoziiert waren. Insbesondere war die Studie daran interessiert zu prüfen, inwieweit die Bedeutung individueller Lernvoraussetzung unter Berücksichtigung von lehrkraftseitigen Merkmalen zur Vorhersage motivationaler Zielkriterien variierte.

Fragestellung

Folgende Forschungsfragen wurden konkret untersucht:

1) Welche Ausprägungen zeigen sich für die Lehrkräftemerkmale (Lehrziele, Zusammenarbeit mit den Eltern, digitale Kompetenzen) sowie für den individuellen Lernhintergrund von Schülerinnen und Schülern (elterliche Unterstützung, technische Ausstattung zu Hause und Gewissenhaftigkeit)?

2) Welche Zusammenhänge zeigen sich zwischen den Lehrkräftemerkmalen sowie den individuellen und familiären Merkmalen der Schülerinnen und Schüler mit den Zielkriterien Kompetenzerleben, Anstrengungsbereitschaft und Lernfreude?

3) Inwieweit interagieren die unterschiedlichen Merkmale zur Vorhersage der Zielkriterien?

Methode

Zur Beantwortung dieser Fragestellungen wurden Daten einer groß angelegten Onlinestudie genutzt, die im Frühsommer 2020 durchgeführt wurde. Insgesamt nahmen 3.159 Schülerinnen und Schülern, 1.688 Eltern und 227 Lehrkräfte an der Untersuchung teil und wurden jeweils zu den in der Studie betrachteten Einflussgrößen befragt. Für die hierfür verwendeten Skalen zeigten sich zufriedenstellende Reliabilitäten ($.72 \leq \alpha \leq .89$). Lediglich die Skala zur Erfassung sozialer Lehrziele

zeigte eine geringere interne Konsistenz (α = .65). Die Analyse der Daten erfolgte anhand eines multivariaten Mehrebenenmodells.

Ergebnisse

Die Ergebnisse zeigten für die untersuchten Einflussgrößen insgesamt hohe Ausprägungen. Zudem ergaben sich für alle Zielkriterien Klassenunterschiede (.04 ≤ ICC ≤ .13), wenngleich diese Unterschiede weniger deutlich ausgeprägt waren als das aus Studienergebnisse für den Präsenzunterricht bekannt ist. Demgegenüber fielen die Unterschiede innerhalb von Klassen weit höher aus. Insofern betätigten die Ergebnisse die höhere Bedeutsamkeit des individuellen Lernhintergrunds von Schülerinnen und Schülern. Mit Blick auf die konkreten Merkmale fanden sich für die elterliche Unterstützung konsistente Zusammenhänge mit den Zielkriterien (0.08 < β ≤ 0.21, ps < .05), während die Gewissenhaftigkeit der Schülerinnen und Schüler sowie die technische Ausstattung weniger prädiktiv waren (-0.02 < β ≤ 0.12). Darüber hinaus waren die sozialen Lehrziele der Lehrkraft als auch die Zusammenarbeit der Lehrkraft mit den Eltern am vorhersagestärksten (0.13 < β ≤ 0.19, ps < .10). Das Ziel der Gestaltung eines adaptiven Unterrichts oder eines lernförderlichen Unterrichts und auch die durch Lehrkräfte berichteten digitalen Kompetenzen waren hingegen nicht statisch signifikant mit den Zielkriterien assoziiert. Ebenso fanden sich nur einige wenige Interaktionseffekte nach Adjustierung für multiples Testen. Die Interaktionsmuster legten einen kompensatorischen Effekt der lehrerseitigen Merkmale nahe. Die konsistentesten Befunde zeigten sich bezüglich der Zusammenarbeit mit den Eltern für unterschiedliche Lernhintergrundmerkmale und Zielkriterien. Insofern liefern die Studienergebnisse einen wichtigen Ansatzpunkt, um Lernunterschiede zwischen Schülerinnen und Schülern auch in Zeiten des Präsenzunterrichts zu verringern.

A-12: Motivation im Unterrichtskontext

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* T03 R02 D26

Motivation im Unterrichtskontext

Chair(s): **Rebecca Lazarides** (Universität Potsdam, Germany), **Katharina Hettinger** (Universität Potsdam)
*Diskutant*in(nen):* **Markus Dresel** (Universität Augsburg)

Die Motivation Lernender ist von maßgeblicher Bedeutung für Lernerfolg und Lernfreude (Kriegbaum et al., 2018; Loderer et al., 2021) und ist damit eine wichtige Voraussetzung für den Leistungserwerb. Zudem gilt die Motivation Lernender als bedeutsame Zielgröße schulischer Lernprozesse. Im Sinne mehrdimensionaler Bildungsziele dient schulisches Lernen nicht nur der Entwicklung kognitiver Fertigkeiten sondern auch der Entwicklung der Bereitschaft, sich eigeninitiativ mit Inhalten auseinanderzusetzen (Schiepe-Tiska, 2019). Eine motivationsförderliche Unterrichtsgestaltung und eine hohe unterrichtsbezogene Motivation von Lehrkräften sind dabei von Bedeutung für die motivationale Entwicklung Lernender (Lazarides & Schiefele, 2021). Trotz einer Vielzahl an Studien zur motivationalen Entwicklung im Schulverlauf existieren aktuell verschiedene Forschungsdesiderata, die beispielsweise Fragen nach interindividuellen Unterschieden, nach der Bedeutung des Lernkontextes und nach instruktionalen Prozessen betreffen (Eccles & Wigfield, 2020). Das Symposium greift diese Desiderata speziell für motivationale Entwicklungen und Ausprägungen der Motivation im Jugendalter auf und setzt sich vor dem Hintergrund verschiedener Motivationstheorien mit motivationalen Merkmalen Lernender, ihren (interindividuell unterschiedlichen) Entwicklungsverläufen und der Bedeutung des sozialen, fachbezogenen und instruktionalen Kontextes auseinander und fokussiert dabei das Jugendalter, in dem die Motivation Lernender häufig konsistent absinkt (Frenzel et al., 2010). Der erste Beitrag befasst sich vor dem Hintergrund der Erwartungs-Wert Theorie (Eccles & Wigfield, 2020), mit interindividuell unterschiedlichen Veränderungen der intrinsischen Wertüberzeugungen Lernender im Fach Deutsch zu Beginn der Sekundarstufe und untersucht die Bedeutung individueller Merkmale wie Geschlecht, sozioökonomischer Herkunft und der Leistung für die identifizierten Entwicklungsverläufe. Der zweite Beitrag untersucht ausgehend vom Internal/External Frame of Reference (I/E)-Modell (Marsh, 1986) die Bedeutung von sozialen und dimensionalen Vergleichen für Fähigkeitsselbstkonzepte und Wertüberzeugungen von Sekundarschüler:innen in Mathematik und Deutsch und klassenstufenbezogene Unterschiede. Der dritte Beitrag geht bezugnehmend auf die sozialkognitive Theorie (Bandura, 1986) der Frage nach, welche Rolle verschiedene Facetten der Lehrkräfteselbstwirksamkeit für die Veränderungen von Schüler:innenlernfreude und Selbstwirksamkeit Lernender in der Sekundarstufe haben und inwiefern diese Zusammenhänge durch schüler:innenperzipiertes und lehrkräfteperzipiertes Unterrichtshandeln mediiert sind. Der vierte Beitrag untersucht ausgehend von Überlegungen der Kontroll-Wert Theorie (Pekrun, 2016) Veränderungen von Freude und Anstrengungsbereitschaft im Laufe eines Schuljahres in Abhängigkeit von Gestaltungselementen der Lernumgebung – hier regelmäßigen Intervallen selbstgesteuerten Lernens in der Sekundarstufe. Abschließend werden in der Diskussion die vorgestellten Beiträge kritisch beleuchtet und die Frage thematisiert, inwieweit es gelingen kann, die verschiedenen Perspektiven zusammenzuführen.

Die heterogene Entwicklung von Motivation in der Sekundarstufe

Swantje Bolli[1], Andrea Westphal[2], Rebecca Lazarides[1], Miriam Vock[1]
[1]Universität Potsdam, [2]Universität Greifswald

Theoretischer Hintergrund

Über die Schulzeit hinweg nimmt Motivation im Allgemeinen ab (Scherrer & Preckel, 2019), obwohl sie für das Lernen eine wichtige Rolle spielt. Dieser generelle Entwicklungsverlauf verdeckt jedoch potenzielle interindividuelle Unterschiede. Aktuelle Erkenntnisse deuten darauf hin, dass es durchaus Gruppen von Schüler:innen gibt, deren Motivation stabil bleibt oder sich positiv entwickelt und dass sich diese Gruppen hinsichtlich ihrer Ausgangsmotivation und ihres Entwicklungsverlaufes unterscheiden lassen (Archambault et al., 2010; Gaspard et al., 2020; Wang et al., 2017). Weshalb sich die Motivation unterschiedlich entwickelt ist noch nicht abschließend erklärt. Zum einen wurden hohe Leistungen in den unteren Klassenstufen als ein Faktor genannt, der zu einem höheren Motivationsverlauf beiträgt (Archambault et al., 2010; Wang et al., 2017). Zum anderen könnten unterrichtsbezogene Variablen wie die Unterrichtsqualität die unterschiedlichen Verläufe bedingen. In verschiedenen Arbeiten wird daher vorgeschlagen, diese in die Untersuchung der Motivationsverläufe miteinzubeziehen (Gaspard et al., 2020; Musu-Gillette et al., 2015), zumal unterrichtsbezogene Merkmale systematisch angepasst werden könnten, um ein motivierendes Klima zu fördern. In dieser Studie wird vor dem Hintergrund des Erwartung-Wert-Modells von Eccles et al. (1983) die interindividuelle Veränderung des intrinsische Wertes Lernender untersucht.

Fragestellung

Um die Motivation einzelner Lernender effektiv fördern zu können, ist es notwendig interindividuelle Unterschiede sowie ihre Entstehungsbedingungen und Konsequenzen besser zu verstehen. Das Ziel dieser längsschnittlichen Studie ist es daher, interindividuell unterschiedliche motivationale Entwicklungen von Schüler:innen im Verlauf der Sekundarstufe und relevante Faktoren für diese Entwicklungsverläufe zu untersuchen. Dabei werden folgende Fragen untersucht:

1. Welche Gruppen von Schüler:innen gibt es, die sich systematisch in ihrem motivationalen Entwicklungsverlauf unterscheiden?

2. Welche Rolle spielen Leistung und Unterrichtsqualität für interindividuell unterschiedliche Entwicklungsverläufe?

Methode

Die analysierten Daten stammen aus der BiKS-8-14 Studie (Artelt et al., 2013). Insgesamt wurden N = 1405 Schüler:innen des Gymnasiums von Klassenstufe 5-7 untersucht (t1: 47% m; Malter = 11.08, SDalter = 0.39). Um die motivationale Entwicklung zu verfolgen, wurde der intrinsische Wert für die Domäne Deutsch jährlich mittels zwei Items auf einer 5-stufigen Likert-Skala erhoben. Um die Heterogenität von Motivationsverläufen zu untersuchen, wurden Growth Mixture Modelle entsprechend den Protokollen von Soloski und Durtschi (2020) und van der Nest et al. (2020) spezifiziert. Zunächst wurde die latente Wachstumskurve für die gesamte Stichprobe identifiziert und dann die Anzahl der distinkten Verläufe bestimmt. Als Nächstes wurde die Ausgangsmotivation sowie das Ausmaß der motivationalen Veränderung zwischen den Verläufen verglichen. Anschließend wird

die Leistung sowie die schülerperzipierte Unterrichtsqualität in das Modell mit einbezogen, um deren Auswirkung auf die Verlaufszugehörigkeit und Wachstumsfaktoren zu prüfen.

Ergebnisse und Bedeutung

Die Ergebnisse der Growth Mixture Analysen deuten anhand zufriedenstellender Modellparameter darauf hin, dass es fünf unterscheidbare lineare Motivationsverläufe für den intrinsischen Wert von Deutsch von Klasse 5-7 gibt. Die Mehrheit der Schüler:innen (39%) beginnt die Sekundarstufe mit einem mittelstarken intrinsischen Wert, der dann abnimmt. Zudem gibt es viele Schüler:innen (36%), die mit schwach ausgeprägter Motivation starten und ihre Motivation nur leicht steigern können. Einige Schüler:innen schaffen es, ihre zunächst niedrige Motivation stark zu erhöhen (5%), wohingegen andere bereits mit hoher Motivation starten und diese dann leicht (9%) oder sehr stark (10%) abfällt. Die Ergebnisse stimmen damit mit bisherigen Forschungsbefunden überein, bei denen 3-7 distinkte Motivationsverläufe identifiziert wurden (Archambault et al., 2010; Guo et al., 2018). Ergebnisse zu Fragestellung 2 stehen noch aus. Hier weist die Literatur auf leistungs- und geschlechtsbezogene Unterschiede sowie auf relevante Unterrichtsmerkmale hin (Gaspard et al., 2020). Die Ergebnisse verdeutlichen, dass interindividuell unterschiedliche Verläufe identifiziert werden können.

Vergleichseffekte der Leistungen von Schüler*innen auf ihre Fähigkeitsselbstkonzepte und Wertüberzeugungen in der Sekundarschulzeit

Hanna Gaspard[1], Ann-Kathrin Jaekel[2], Richard Göllner[2]
[1]Techische Universität Dortmund, [2]Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

Theoretischer Hintergrund

Ausgehend vom Internal/External Frame of Reference (I/E)-Modell (Marsh, 1986) nutzen Schüler*innen zwei Bezugsrahmen zum Vergleich ihrer Leistungen: (a) soziale Vergleiche, wobei sie ihre Leistungen mit denen ihrer Mitschüler*innen vergleichen (externer Bezugsrahmen) und (b) dimensionale Vergleiche, wobei sie ihre Leistungen mit ihren eigenen Leistungen in anderen Fächern vergleichen (interner Bezugsrahmen). Die Forschung hat gezeigt, dass solche Vergleiche das Fähigkeitsselbstkonzept (Möller et al., 2020), aber auch die Wertüberzeugungen (Gaspard et al., 2018) beeinflussen. Solche Vergleichseffekte fallen typischerweise stärker aus, wenn Noten im Vergleich zu standardisierten Tests genutzt werden (Möller et al., 2020). Nach bisheriger Forschung scheinen solche Vergleichseffekte zudem für ältere Schüler*innen stärker auszufallen (Möller et al., 2020). Allerdings fehlt es an Studien, in denen anhand ausreichend großer Stichproben systematisch untersucht wird, ob solche Vergleichseffekte für Wertüberzeugungen ebenfalls vom Alter abhängen (siehe Gaspard et al., 2018).

Fragestellung

In der vorliegenden Studie wollen wir daher untersuchen, (a) in welchem Ausmaß Fähigkeitsselbstkonzepte und Wertüberzeugungen von Sekundarschüler*innen in Mathematik und Deutsch durch soziale und dimensionale Vergleiche ihrer Noten und Testleistungen beeinflusst werden und (b) in welchem Ausmaß solche Vergleichseffekte in Abhängigkeit der Klassenstufe variieren.

Methode

Daten wurden an 11 Gymnasien, 8 Realschulen, 1 Werkrealschule und 7 Gemeinschaftsschulen in Baden-Württemberg zu zwei Zeitpunkten im Abstand eines Schuljahres erhoben. Für die vorliegenden Analysen wurden Daten von 6.479 Schüler*innen der Klassenstufen 5-10 zum ersten Messzeitpunkt genutzt. Das Fähigkeitsselbstkonzept wurde in Mathematik und Deutsch mit je vier Items erfasst (α = .88/.93). Die Wertüberzeugungen in beiden Fächern wurden mit je drei Items pro Wertkomponente (intrinsischer Wert, Wichtigkeit, Nützlichkeit und Kosten) erfasst (α = .81 - .96). Für die motivationalen Skalen konnte starke Messinvarianz über die Klassenstufen bestätigt werden. Die Schüler*innen berichteten ihre eigenen Noten. Zudem bearbeiteten sie zwei kurze Tests zur Erfassung ihrer Basiskompetenzen in Mathematik (MBK 5-12+; Ennemoser et al., 2011; Krajewski & Ennemoser, 2013) und Lesen (LGVT 5-12+; Schneider et al., 2017), welche zufriedenstellende Reliabilitäten über Schulformen und Jahrgangsstufen hinweg zeigten (LGVT 5–12+: r = .93; MBK 5–12+: r = .79). In separaten Strukturgleichungsmodellen wurden die Fähigkeitsselbstkonzepte und Wertüberzeugungen in Mathematik und Deutsch auf die Leistungsindikatoren in beiden Fächern und die Klassenstufe regrediert. Zusätzlich wurden Interaktionen zwischen den Leistungsindikatoren und der Klassenstufe aufgenommen, um Unterschiede in den Vergleichseffekten in Abhängigkeit der Klassenstufe zu testen.

Ergebnisse & Bedeutung

Im Einklang mit dem I/E-Modell zeigten sich signifikante Pfade von den Noten und Testleistungen auf das Fähigkeitsselbstkonzept und Wertüberzeugungen im entsprechenden und nicht-entsprechenden Fach. Hier gab es wenige Ausnahmen: der Effekt des Lesetests auf Wichtigkeit in Deutsch und die Effekte der Testleistungen auf Kosten im nicht-entsprechenden Fach waren nicht signifikant. Fähigkeitsselbstkonzepte und Wertüberzeugungen werden also durch soziale und dimensionale Vergleiche beeinflusst. Diese Pfade fielen wie in früherer Forschung stärker für die Noten aus als für die Testleistungen. Mit wenigen Ausnahmen sagte die Klassenstufe Fähigkeitsselbstkonzept und (positive) Wertüberzeugungen negativ und Kosten positiv vorher. Dies entspricht den Ergebnissen früherer Forschung zu einer Abnahme von Fähigkeitsselbstkonzept und Wertüberzeugungen über die Sekundarschulzeit. Schließlich fanden wir in Bezug auf die Interaktionsterme, dass 15 von 20 Pfaden innerhalb der Fächer (im Sinne sozialer Vergleiche) und 7 von 20 Pfaden zwischen Fächern (im Sinne dimensionaler Vergleiche) signifikant von der Klassenstufe abhingen. Mit einer Ausnahme wiesen diese Interaktionsterme auf stärkere Vergleichseffekte für Schüler*innen in höheren Klassenstufen hin. Diese stärkere Bedeutung von Vergleichseffekten mit zunehmendem Alter kann dazu beitragen zu erklären, warum die Zusammenhänge zwischen Leistungen und motivationalen Überzeugungen innerhalb von Fächern mit dem Alter stärker werden (Denissen et al., 2007) und warum die Zusammenhänge motivationaler Überzeugungen zwischen Fächern mit dem Alter schwächer werden (Wan et al., 2021). Darüber hinaus ist geplant, zur Konferenz Analysen für Vergleichseffekte auf motivationale Veränderung zum zweiten Testzeitpunkt zu berichten.

Die Bedeutung spezifischer Lehrkräfteselbstwirksamkeit für Freude und Selbstwirksamkeit von Schüler:innen im Mathematikunterricht - Unterrichtsverhalten als vermittelndes Merkmal

Katharina Hettinger, Rebecca Lazarides, Ulrich Schiefele
Universität Potsdam

Theoretischer Hintergrund

Die Selbstwirksamkeit von Lehrkräften ist theoretisch bedeutsam für deren Unterrichtshandeln (Bandura, 1997; Tschannen-Moran et al., 1998) und ist ein wesentliches Merkmal von Lehrkräftemotivation mit hoher Relevanz für affektiv-motivationale Merkmale von Lernenden (Lauermann & ten Hagen, 2021; Zee & Koomen, 2016). Unklar ist jedoch, inwiefern Selbstwirksamkeit von Lehrkräften über das Unterrichtshandeln mit der Motivation von Lernenden längsschnittlich zusammenhängt. Bisher existiert meist nur querschnittliche Forschung zu den Zusammenhängen von Lehrkräftemotivation hin zu Lernendenmotivation über Unterrichtsverhalten (z.B., Burić & Kim, 2020). Die wenigen vorhandenen längsschnittlichen Studien zeigen inkonsistente Befunde: Manche Studien finden signifikant positive Zusammenhänge zwischen Lehrkräfteselbstwirksamkeit und Lernendenmotivation, die über schülerperzipierte Unterrichtsmerkmale vermittelt werden (Oppermann & Lazarides, 2021), während andere Studien keine Zusammenhänge zwischen Lehrkräfteselbstwirksamkeit und schülerperzipierten Unterrichtsmerkmalen finden (Holzberger et al., 2013; Praetorius et al., 2017). Zudem untersuchen wenige Studien wie spezifische Facetten der Lehrkräfteselbstwirksamkeit über Unterrichtshandeln auf motivational-affektive Merkmale Lernender wirken. Aktuelle Forschungsarbeiten verweisen jedoch auf die Notwendigkeit eines „matching“ von inhaltlich verwandten Konstrukten bei der Betrachtung von Lehrkräftemotivation und deren Unterrichtsverhalten (Bardach & Klassen, 2021), um die Zusammenhänge zwischen Lehrkräftemotivation und Motivation Lernender besser zu verstehen.

Fragestellung

Vor diesem Hintergrund untersucht die vorgestellte Studie längsschnittliche Zusammenhänge zwischen fünf spezifischen Selbstwirksamkeitsdimensionen von Lehrkräften und der Selbstwirksamkeit und Freude von Schüler:innen in Mathematik über wahrgenommenen Unterricht. Dabei untersucht die vorliegende Studie, welche Zusammenhänge zwischen a) spezifischen Selbstwirksamkeitsfacetten von Lehrkräften und (von Schüler:innen und Lehrkräften) wahrgenommenem Unterricht, b) von Schüler:innen und Lehrkräften wahrgenommenem Unterricht und der Selbstwirksamkeit sowie Freude der Schüler:innen und c) indirekten Zusammenhänge zwischen spezifischen Selbstwirksamkeitsfacetten von Lehrkräften und der Selbstwirksamkeit sowie Freude der Schüler:innen über wahrgenommen Unterricht bestehen.

Methode

Basierend auf längsschnittlichen Daten einer Large-Scale Studie, wurden 50 Mathematiklehrkräfte (66.0% weiblich; 94.0% in Deutschland geboren) und deren 959 Schüler:innen (48.7% weiblich; MAlter = 14.20, SD = 0.62; Range: 13-17 Jahre; 91.7% in Deutschland geboren) aus 30 weiterführenden Schulen (52 Klassen) in Berlin und Brandenburg in die Analysen eingeschlossen. Zu drei Messzeitpunkten, T1: Beginn der neunten Jahrgangsstufe; T2: Beginn der zehnten Jahrgangsstufe; T3: Mitte der zehnten Jahrgangsstufe, füllten die Schüler:innen in zwei Terminen Fragebögen aus und absolvierten einen standardisierten Mathematikkompetenztest, während deren Lehrkräfte ebenfalls (zeitgleich mit den Schüler:innen) Fragebögen ausfüllten und einen bildungswissenschaftlichen Kompetenztest absolvierten. Zur Beantwortung der Forschungsfragen

wurden latent-manifeste Mehrebenenanalysen (Marsh et al., 2009) sowie cross-level Mediationsmodelle (Pituch & Stapleton, 2012) in Mplus berechnet.

Ergebnisse und ihre Bedeutung

Die Analysen zeigten, dass Selbstwirksamkeit für Klassenmanagement (T1) positiv mit auf Klassenebene erlebter Freude (T3) über schülerwahrgenommene Disziplin (T2) zusammenhängt (unique class-level indirect effect). Zudem zeigte sich, dass die Selbstwirksamkeit für Schülerengagement (T1) positiv mit individueller Selbstwirksamkeit Lernender (T3) über schülerwahrgenommene Kompetenzunterstützung (T2) – sowie, dass die Selbstwirksamkeit für schülerorientiertes Lehren (T1) positiv mit individueller Selbstwirksamkeit Lernender (T3) über schülerwahrgenommene lernzielorienterte Instruktion (T2) zusammenhängt (cross-level indirect effects). In allen Analysen wurde auf den Ausgangswert der Freude bzw. der Selbstwirksamkeit der Schüler:innen kontrolliert. Die Ergebnisse verweisen demnach auf eine hohe Bedeutsamkeit von Selbstwirksamkeit für Klassenmanagement für die in der Klasse erlebte Freude – als geteilter Affekt. Wohingegen Lehrkräfteselbstwirksamkeitsfacetten mit Fokus auf motivationale Prozesse (für Schülerengagement und für schülerorientiertes Lehren) bedeutsam für die individuelle Selbstwirksamkeit von Schüler:innen sind. Die Ergebnisse dieser Studie tragen zur Theorieentwicklung bei, indem spezifische Pfade der Lehrkräfteselbstwirksamkeit auf Unterrichtshandeln und motivationale sowie affektive Merkmale von Schüler:innen identifiziert werden und dadurch die unterschiedliche Bedeutsamkeit von Lehrkräfteselbstwirksamkeit für individuelle oder als Gruppe erlebte Motivation und Emotionen von Schüler:innen herausgestellt wird. Weiterhin scheint die Selbstwirksamkeit von Lehrkräften indirekt über schülerperzipierten Unterricht auf Freude sowie Selbstwirksamkeit der Schüler:innen übertragen zu werden, wobei die von Schüler:innen wahrgenommenen Unterrichtsaspekte vermitteln. Praktische Implikationen beziehen sich auf die gezielte Förderung von Lehrkräfteselbstwirksamkeit in schulischen Kontexten.

Freude und Anstrengung von adoleszenten Schüler*innen im Unterricht: Beeinflusst durch Intervalle selbstgesteuerten Lernens

Diana Raufelder[1], Sabine Schweder[2]
[1]Ernst-Moritz Arndt University Greifswald, [2]Ernst-Moritz Arndt Universität Greifswald

Theoretischer Hintergrund

Was ermutigt Schüler*innen, Spaß an Lernaufgaben zu haben und sich dafür anzustrengen? Dies ist eine kritische Frage für Bildungsprozesse, denn das Ausmaß an Freude und Anstrengung beeinflusst nicht nur das Lernen der Schüler*innen, sondern auch deren Leistungsniveau (Linnenbrink-Garcia, Patall, & Pekrun, 2016; Linnenbrink-Garcia & Pekrun, 2011). Die Kontroll-Wert-Theorie (Pekrun, Elliot, & Maier, 2006) besagt, dass Leistung, Emotionen und motivationale Merkmale „durch Merkmale der Lernumgebung, einschließlich des motivationalen Milieus des Klassenzimmers (z. B. Zielstrukturen und Autonomieunterstützung) und der kognitiven Qualität der Lernaufgaben (z. B. optimales Niveau der Aufgabenherausforderung), geformt werden können“ (Putwain et al., 2021, S. 349). In Anbetracht dieser Tatsachen ist es wichtig, Erkenntnisse über die wechselseitigen Zusammenhänge zwischen Freude, Anstrengung und Unterrichtsmerkmalen zu gewinnen. Der untersuchte Unterrichtsansatz beinhaltet zweimal im Schuljahr Intervalle für selbstgesteuertes Lernen (SDL). Schulen, die nach

effektiveren Unterrichtsmethoden suchen, versuchen, den Unterricht durch SDL-Intervalle zu erweitern. Während dieser Phasen legen die Lehrer*innen die Kontrolle über das Lernen in die Hände der Schüler*innen, die zu Verantwortlichen für ihre eigenen Lernbedingungen werden.

Fragestellung

In der vorliegenden Studie sollte überprüft werden, ob SDL die Entwicklung von Freude und Anstrengung im Laufe der Zeit - in diesem Fall während eines einzigen Schuljahres - begünstigt. Um die möglichen Auswirkungen von Veränderungen in der Lernumgebung direkt beobachten zu können, wurde SDL als einwöchige Intervention zweimal im Jahr in den regulären, von Lehrer*innen geleiteten Unterricht integriert. Die folgenden Forschungsfragen wurden untersucht: (1) Welche Rolle spielen die spezifischen Qualitäten der Lernumgebung für die allgemein positive Entwicklung von Freude und Anstrengung im Verlauf eines Schuljahres? (2) Hängen Freude und Anstrengung nicht nur innerhalb eines einzelnen Messzeitpunkts zusammen, sondern auch im Zeitverlauf?

Methode

Auf der Grundlage der Kontroll-Wert-Theorie untersuchte die vorliegende Studie die Entwicklung und Veränderung von Freude und Anstrengung während eines Schuljahres. Sie analysierte Fragebogendaten von jugendlichen Schüler*innen (N = 754; MAge = 13,56; SD = 1,2; 49,7% weiblich), die zweimal an einer einwöchigen Intervention zum selbstgesteuerten Lernen teilnahmen. Die Schüler*innen stammten aus 37 Klassen der 6. bis 9. Jahrgangsstufe an sechs ländlichen, gemischt gegliederten Schulen in Mecklenburg-Vorpommern und Schleswig-Holstein, Deutschland. Diese Schulen haben ab der 6. Klasse systematisch Intervalle für selbstgesteuertes Lernen in den Lehrplan aufgenommen. In jedem Schulhalbjahr haben die Schüler*innen eine Schulwoche Zeit, um einer Frage nachzugehen, die ausschließlich auf dem Interesse der Schüler*innen beruht. Diese Lehrplankonfiguration wird in den an der Studie beteiligten Schulen bis zur 9. Klasse fortgesetzt; die Schüler*innen nehmen also an bis zu acht Intervallen für selbstgesteuertes Lernen (SDL) teil. Die Stichprobe der vorliegenden Studie beschränkt sich auf Schüler*innen, die zum ersten Mal an der SDL-Konfiguration teilgenommen haben. Mithilfe von Mplus wurde ein bivariates latentes neighbor change Modell durchgeführt, um die Assoziationen zwischen den Δ beider Variablen innerhalb der Zeit und im Zeitverlauf zu untersuchen.

Ergebnisse und ihre Bedeutung

Die Ergebnisse des bivariaten latenten neighbor change Modell zeigen, dass - im Gegensatz zu früheren Studienergebnissen - im Laufe des Schuljahres generell eine positive Entwicklung von Freude und Anstrengung zu verzeichnen war und dass der Anstieg hauptsächlich auf die beiden Interventionen des selbstgesteuerten Lernens zurückzuführen war. Darüber hinaus zeigen die Ergebnisse, dass Freude und Anstrengung im Laufe der Zeit wechselseitig miteinander verbunden waren, allerdings nur dann, wenn zuerst selbstgesteuertes Lernen praktiziert wurde. Durch die Intervalle des selbstgesteuerten Lernens scheint es möglich, dem Absinken von Freude und Anstrengung entgegenzuwirken. Dieses Absinken ist besonders ausgeprägt beim Eintritt in die Sekundarstufe und zu Beginn der Adoleszenz.

A-13: Unterricht während der Pandemie: Was wissen wir über die Qualität des Distanzunterrichts?

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* T03 R02 D39

Unterricht während der Pandemie: Was wissen wir über die Qualität des Distanzunterrichts?

Chair(s): **Verena Jörg** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main)
*Diskutant*in(nen):* **Christoph Helm** (PH Zug)

Die Schulschließungen im Zuge der Covid-19-Pandemie veränderten die Arbeitsbedingungen von Lehrkräften grundlegend. Studien zeigen, dass viele Lehrkräfte die Anpassung an das Unterrichten über Distanz als große Schwierigkeit wahrnahmen und von einer hohen emotionalen Erschöpfung berichteten (z.B. Kim et al., 2022; Pellerone, 2021). Befragungen wiesen zudem darauf hin, dass die Art und Häufigkeit der eingesetzten Unterrichtsformen über Distanz zwischen Lehrkräften stark variierte (Helm et al., 2021). Jedoch gibt es bislang kaum Studien zur Qualität des Unterrichts während der Pandemie.

Ziel des Symposiums ist es daher, Forschungsarbeiten zur Qualität des Distanzunterrichts zu bündeln. Hierbei werden gezielt verschiedene Disziplinen wie die Fachdidaktik, pädagogische Psychologie und Erziehungswissenschaft berücksichtigt und Studien mit verschiedenen Methoden zusammengebracht. Die Bandbreite reicht von der Analyse authentischer Unterrichtsmaterialen aus der Zeit der Schulschließungen über qualitative Interviews mit Lehrkräften bis hin zur Analyse von Einschätzungen der Unterrichtsqualität aus Sicht der Lehrkräfte und Schüler:innen.

Der erste Beitrag beleuchtet die Qualität der kognitiven Aktivierung im Mathematik-Unterricht während der ersten Schulschließungen im Frühjahr 2020. In der explorativen Studie wurden mehrere hundert Materialien anhand mathematik-didaktischer Kriterien klassifiziert. Untersucht wurde das Potenzial zur kognitiven Aktivierung innerhalb der Lerninhalte sowie die Qualitätsentwicklung innerhalb des 12-wöchigen Beobachtungszeitraums.

Der zweite Beitrag geht der Frage nach, ob sich die Qualität des Mathematikunterrichts über Distanz anhand der für den Präsenzunterricht vielfach herangezogenen Basisdimensionen der Unterrichtsqualität (Potenzial zur kognitiven Aktivierung, konstruktive Unterstützung und Klassenführung/Strukturierung) systematisch beschreiben lässt. Zudem werden Zusammenhänge der Qualität der Unterrichtsmaterialien zu potenziellen Ressourcen (z.B. technologisch-pädagogisches Wissen der Lehrkräfte oder technische Ausstattung der Schulen) analysiert.

Der dritte Beitrag untersucht, inwiefern Schüler:innen verschiedene Dimensionen der Unterrichtsqualität über mehrere Fächer hinweg unterschiedlich beurteilen und wie diese Beurteilungen mit der Motivation der Schüler:innen zusammenhängen. Zudem wird die Qualität des Distanzunterrichts mit der Qualität des Präsenzunterrichts verglichen. Dazu wurden mehrere tausend Schüler:innen der Sekundarstufe I im Sommer 2019 und im Frühsommer 2021 zur Unterrichtsqualität im Deutsch- und Mathematik-Unterricht befragt.

Im vierten und letzten Beitrag steht die Qualität der Beziehung zwischen Lehrkräften und Schüler:innen in der Zeit der Schulschließungen im Fokus. Es wird anhand von Daten qualitativer Interviews mit Lehrkräften untersucht, welche Faktoren zu einer positiven Beziehung zwischen

Lehrer:innen und Schüler:innen beitragen und wie die Qualität der Beziehung mit Merkmalen der Unterrichtsqualität zusammenhängt.

Die Bündelung dieser Beiträge mit unterschiedlichen methodischen Zugängen trägt zu einem besseren Verständnis der Qualität des Unterrichts während der Schulschließungen bei. Zudem bilden die Erkenntnisse der Studien die Grundlage, um zu diskutieren, welche Implikationen sich für den Präsenzunterricht ableiten lassen.

Beiträge des Symposiums

Fernunterricht während des ersten Lockdowns – eine explorative Studie zur kognitiven Aktivierung durch Aufgaben im Mathematikunterricht der Sekundarstufe

Bruno Scheja[1], **Benjamin Rott**[2]
[1]St.-Ursula-Gymnasium Brühl, [2]Universität zu Köln

Die Corona-Pandemie stellte die Bildungsbeteiligten vor Herausforderungen bisher unbekannten Ausmaßes. Dies zeigt sich z. B. an der Notwendigkeit der Durchführung eines zum Teil mehrere Monate andauernden Fernunterrichts – zunächst annähernd ohne Nutzung unterstützender digitaler Lehr-Lehr-Formate. Dabei bleibt bisher im Dunkeln, wie kognitiv gehaltvoll Lehrkräfte unter solchen Bedingungen ihren Fachunterricht gestaltet haben. Die vorliegende explorative Studie (vgl. Scheja & Rott, 2022) geht auf Grundlage zentraler Kriterien zur Beurteilung der Qualität des Fernunterrichts diesem Fragekomplex für das Fach Mathematik nach.

Dazu wird kognitive Aktivierung als eine der drei Qualitätsdimensionen der Tiefenstrukturen im Zusammenhang mit ihrem wichtigsten „Träger“, den Mathematikaufgaben (s. Baumert & Kunter, 2013; Hiebert et al., 2003), betrachtet. Die Aufgaben, die im Mathematikunterricht während des Fernunterrichts eingesetzt wurden, werden so klassifiziert, dass damit ihr Beitrag zur kognitiven Aktivierung der Lerngelegenheiten identifizierbar wird. Dazu wird eine Auswahl empirisch bewährter Klassifikationskategorien genutzt, die der Mathematikdidaktik sowie auch der allgemeinen Didaktik entstammen. Dies sind: (I.) Typ mathematischen Arbeitens, (II.) Argumentieren, (III.) kognitive Prozesse, (IV.) Repräsentationsformen des Wissens und (V.) Anzahl der geforderten Lösungswege. Unter anderem an den Lernzieltaxonomien von Anderson und Krathwohl (2001) und dem Klassifikationssystem des COACTIV-Projekts (Neubrand et al., 2011) orientierend, werden zu jeder Kategorie drei Kategorie-Ausprägungen festgelegt, so dass der kognitive Anspruch in Form spezifischer Tätigkeiten identifizierbar wird und von Ausprägung zu Ausprägung ansteigt; jede der fünf Kategorien kann somit als ein Anforderungsmerkmal des Lösungsprozesses betrachtet werden, das seinerseits im Hinblick auf die kognitive Komplexität einstufbar ist.

Die Untersuchung des Aufgabensatzes (N=457) – einer lokalen Konvenienzstichprobe – von acht Lehrenden der Sekundarstufe I (sechs Gymnasial- und zwei Realschullehrkräfte) erfolgt unter den Perspektiven a) der kognitiven Aktivierung (Untersuchungszeitraum: erster Lockdown zwischen März und Juni 2020 (12 Schulwochen)) und b) der zugehörigen zeitlichen Entwicklung (Zeitraum 1: die ersten drei Wochen des Fernunterrichts; Zeitraum 2: die letzten drei Wochen des Fernunterrichts).

Die Befunde weisen a) eher auf ein sehr geringes durch die Aufgaben vermitteltes Aktivierungspotential während des Corona-Zeitraums hin. Zu dieser Charakterisierung trägt insbesondere bei, dass kategorienübergreifend kognitiv aktivierungsarme Kategorienausprägungen

festgestellt wurden. Dies wird darüber hinaus nicht nur am weitgehenden Fehlen von Aufgaben mit Aktivitäten der mittleren und vor allem auch hohen Ausprägung deutlich, die erst eine aktive Verknüpfung der Lern-Gegenstände anregen würden, sondern auch an der Berücksichtigung zahlreicher gleichartiger Aufgaben, bei denen das Umgehen mit Kalkülen (kontextfrei oder eingekleidet) eine zentrale Anforderung darstellt. Dieses Profil liefert durch ein weitgehendes Fehlen von Exploration, Scaffolding, Argumentationen und Artikulation (Helm et al., 2021) zugleich erste empirische Hinweise auf eine Bestätigung der Vermutung, dass sich kognitive Aktivierung im Fernunterricht besonders schlecht garantieren lässt (Klieme, 2020, S. 127).

Unter b) der Perspektive einer zeitlichen Entwicklung zeigt sich dieses Profil zudem eher stabil und weist allenfalls punktuelle Entwicklungen auf. So ist sowohl die Art des verlangten Wissens, die kleine Spannweite der in den Aufgaben verankerten Anforderungen sowie auch die niedrige kognitive Komplexität dieser innerhalb der beiden Zeiträume miteinander vergleichbar.

Distanzunterricht in der Pandemie - Analyse der Unterrichtsqualität anhand von Unterrichtsmaterialien

Nikolaus Bönke[1], Verena Jörg[2], Patrick Schreyer[3], Thamar Voss[1], Uta Klusmann[4], Mareike Kunter[2]
[1]Universität Freiburg, [2]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main, [3]Universität Kassel, [4]IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel

Die Schulschließungen während der Covid-19-Pandemie führten zu einem stark veränderten Unterrichtsgeschehen. Während zahlreiche Forschungsarbeiten die Auswirkung des Distanzunterrichts auf das Wohlbefinden von Schüler:innen, Erziehungsberechtigten oder Lehrkräften untersuchten, blieben Arbeiten, die sich mit der Qualität des Distanzunterrichts beschäftigen, rar (Helm et al., 2021). Einige wenige Selbstberichtsstudien weisen auf substanzielle Qualitätsunterschiede in Abhängigkeit der eingesetzten Lernformen und Methoden hin (z. B. Jaekel et al., 2021). Zudem zeigen Studienergebnisse, dass das technologisch-pädagogische Wissen (TPK) der Lehrkräfte sowie die technische Ausstattung der Schulen wichtig für einen qualitätsvollen Distanzunterricht sind (König et al., 2020).

Ziel dieser Studie ist daher erstens, die Unterrichtsqualität im Distanzunterricht anhand von eingesetzten Unterrichtsmaterialien zu beschreiben. Die theoretische Basis bildet das Modell der drei Basisdimensionen der Unterrichtsqualität, das im deutschen Sprachraum vielfach zur Abbildung von Unterrichtsprozessen genutzt wird (Klieme, 2020). Es wird untersucht, ob das Modell, das auf der Grundlage von Forschung zum Präsenzunterricht entstanden ist, geeignet ist, die Qualität des Distanzunterrichts zu systematisieren. Zweitens wird der Zusammenhang zwischen Unterrichtsqualität und den Ressourcen der Lehrkräfte analysiert.

Die Stichprobe umfasst N=60 Mathematiklehrkräfte der Sekundarstufe I und II (M(SD)Alter=42.21(3.00), 61% weiblich), die im Rahmen einer bestehenden Längsschnittstudie, während der (teilweisen) Schulschließungen im Sommer 2021 an einer Online-Befragung teilnahmen. Sie reichten erstens Materialien ein, die sie im Distanzunterricht zur Einführung eines Themas verwendet haben. Zweitens bearbeiteten die Lehrkräfte Fragebögen zu den Ressourcen Technische Ausstattung der Schule und TPK sowie zur Einschätzung ihres Unterrichts hinsichtlich dessen Strukturierung, konstruktiver Unterstützung und kognitiver Aktivierung.

Zur Bewertung der Qualität der Unterrichtsmaterialien wurde in Anlehnung an Instrumente der COACTIV- (Jordan, 2006) und TALIS-Video-Studie (Schweig & Stecher, 2020) ein Kodier-Schema entwickelt, das die Bewertung der Materialien anhand der eingesetzten Methoden (z.B. Art des Mediums, der Aufgabe) sowie Codes zur Qualität des Unterrichts (d.h. zu Aspekten der Strukturierung, Unterstützung und kognitiven Aktivierung) beinhaltet. Drei geschulte Rater:innen bewerteten die Materialen anhand der 20 Codes des Kodier-Schemas (Ratingskala 1-3). Für alle Codes wurde eine ausreichende bis sehr gute Inter-Rater-Reliabilität (ICC=.65-.99) erreicht.

Die Ratings weisen auf eine hohe Variabilität in den Materialien hin, die von gescannten handschriftlichen Notizen bis hin zu animierten, interaktiven Aufgaben reicht. Die Bewertung der Materialien weist auf ein geringes Ausmaß an Individualisierung und Differenzierung hin, und dass selten zum Austausch zwischen den Schüler:innen angeregt wurde. Positiv fallen der häufige Einsatz von unterstützenden Elementen, Feedback, Echtweltbezügen und Verknüpfungsaufgaben auf.

Bezüglich der Dimensionen von Unterrichtsqualität zeigt eine konfirmatorische Faktorenanalyse, dass die Daten dem dreidimensionalen Basismodell der Unterrichtsqualität nicht gut entsprechen ($\chi 2$=117.87, df=74, p<.01, CFI=.829, RMSEA=.099, SRMR=.110). In einer exploratorischen Faktorenanalyse finden wir Evidenz für ein Fünf-Faktor-Modell mit den Faktoren herauffordernde Aufgaben (M(SD)=1.94(.60)), Unterstützung individueller Aufgabenbearbeitung (M(SD)=1.37(.37)), Lernziele und Feedback (M(SD)=1.85(.48)), Strukturierung (M(SD)=1.40(.45)), und Soziale Eingebundenheit (M(SD)=1.49(.47)). Die Übereinstimmung zwischen den Ratings der Unterrichtsmaterialien und dem Selbstbericht der Lehrkräfte waren gering (p ≥.08).

TPK und die technische Ausstattung der Schulen standen ebenso in keinem signifikanten Zusammenhang mit der Unterrichtsqualität, weder im Selbstbericht noch im Hinblick auf die Unterrichtsmaterialien. Auffällig war, dass die Unterstützung individueller Aufgabenbearbeitung in den Materialien an Gymnasien höher eingeschätzt wurde, als bei nicht-gymnasialen Schulen (t(56)=2.14, p=.04). Im Selbstbericht bewerteten Gymnasiallehrkräfte das Potenzial zur kognitiven Aktivierung als höher (t(56)=2.32, p=.02).

Zusammenfassend ergänzt unsere Studie bestehende Forschungsarbeiten um Daten, die Rückschlüsse auf den Unterricht und dessen Qualität während der Pandemie zulassen. Es zeigt sich, dass es mehr als drei Dimensionen bedarf, um die Komplexität der Situation im Distanzunterricht adäquat abzubilden, und dass Diskrepanzen vorliegen zwischen dem, was der Selbstbericht misst und dem, was die Ratings über die Materialien aussagen (vgl. Clausen, 2000). Das entwickelte Kodier-Schema bildet eine Basis für zukünftige Untersuchungen der Unterrichtsqualität anhand von Materialien. Potentiale, Schwierigkeiten und Ideen für den weiteren Einsatz dieser Bewertungsmethode werden diskutiert.

In Präsenz und auf Distanz: Ein Vergleich von Schülerbeurteilung zur Unterrichtsqualität vor und während der COVID-19-Pandemie

Alicia Protze, Ann-Kathrin Jaekel, Richard Göllner
Universität Tübingen

Die empirische Forschung der vergangenen Jahre zum traditionellen Präsenzunterricht hat gezeigt, dass insbesondere eine unterstützende Beziehung zwischen der Lehrkraft und den Schülerinnen und

Schülern sowie eine strukturierte, konstruktive Lernatmosphäre mit anregenden Aufgaben für den Lernerfolg der Schülerinnen und Schüler bedeutsam sind (Kunter & Trautwein, 2013; Seidel & Shavelson, 2007). Das bis dahin übliche Unterrichtssetting hat sich jedoch durch die Schulschließungen im März 2020 fundamental verändert und Unterricht wurde über viele Wochen hinweg auf Distanz durchgeführt. Auch Forschung aus dem Distanzunterricht zeigt, dass ähnliche Qualitätsmerkmale wie im Präsenzunterricht im Distanzunterricht wichtig für das Lernen der Schülerinnen und Schüler sind (Hawkins et al., 2013; Jaekel et al., 2021). Bislang unklar ist die Frage, ob und inwiefern sich die Wahrnehmung der Unterrichtsqualität aus Sicht der Schülerinnen und Schüler in beiden Unterrichtssettings, dem Präsenz- und dem Distanzunterricht, unterscheiden. Um dies zu adressieren, gehen wir in der vorliegenden Studie den folgenden Forschungsfragen nach:

1) Sind Schülerbeurteilungen der drei Qualitätsdimensionen „Struktur des Unterrichts“, „Unterstützung durch die Lehrkraft“ und „Anspruchsvolle Aufgaben“ für den Deutsch- und Mathematikunterricht zwischen beiden Unterrichtssettings vergleichbar?

2) Inwiefern lassen sich Unterschiede in den Mittelwerten der Qualitätsdimensionen zwischen beiden Unterrichtssettings für beide Fächer finden?

3) Sind die Qualitätsmerkmale beider Fächer mit den Zielkriterien (Anstrengungsbereitschaft und Lernfreude) assoziiert und lassen sich Unterschiede zwischen beiden Un-terrichtssettings finden?

Diese Studie basiert auf Daten zweier groß angelegter Unterrichtsstudien. In der Studie „Unterrichts aus Schülersicht (UNITAS)“, durchgeführt im Sommer 2019, gaben N = 6.298 Schülerinnen und Schüler Urteile zur Unterrichtsqualität im Präsenzunterricht ab. Die Onlinestudie „Lehren und Lernen im Distanzunterricht (CUNITAS)“ wurde kurz nach Ende der ersten Schulschließungsphase im Frühsommer 2020 durchgeführt und basiert auf Angaben von 3.159 Schülerinnen und Schülern. Beide Studien enthalten Angaben zum Deutsch- und Mathematikunterricht in der Sekundarstufe I zu den gleich erfassten Qualitätsdimensionen Struktur, Unterstützung und Anspruchsvolle Aufgaben sowie zu den Zielkriterien Anstrengungsbereitschaft und Lernfreude.

Die Ergebnisse unterschiedlicher Invarianzmodelle zeigten, dass die untersuchten Qualitätsdimensionen beider Fächer in beiden Unterrichtssettings empirisch trennbar und vergleichbar waren. Invarianztests ergaben eine vergleichbare Passung eines mehrdimensionalen Mehrebenenfaktormodells ($\Delta\chi2(-65) = -1.287.00$, $\Delta CFI = .01$, $\Delta TLI = .01$, $\Delta RMSEA = -.00$, $\Delta SRMRw = .00$, $\Delta SRMRb = -.01$). Hinsichtlich der Mittelwertsunterschiede zeigte sich, dass sowohl die wahrgenommene Struktur als auch die Unterstützung im Präsenzunterricht in beiden Fächern positiver eingeschätzt wurde als im Distanzunterricht ($0.28 \leq \Delta \geq 0.37$). Ein gegenteiliges Ergebnis zeigte sich für das wahrgenommene Anspruchsniveau von Aufgaben ($-0.27 \leq \Delta \geq -0.28$). Für die Zielkriterien ergaben sich eine höhere Ausprägung der Lernfreude im Präsenzunterricht ($0.35 \leq \Delta \geq 0.37$), jedoch eine höhere Anstrengungsbereitschaft für den Distanzunterricht ($-0.41 \leq \Delta \geq -0.44$). Beide Zielkriterien waren mit den Unterrichtsbeurteilungen der Schülerinnen und Schüler assoziiert, wenngleich das individuelle Unterrichtserleben in Zeiten des Distanzunterrichts weitaus engere Zusammenhänge mit den Zielkriterien aufwies als für den Präsenzunterricht.

Zusammenfassend zeigten die Ergebnisse, dass der Präsenz- und Distanzunterricht durch Schülerinnen und Schüler entlang ähnlicher Qualitätsdimensionen beurteilt werden kann. Eine besondere Rolle kommt der Unterrichtsqualitätsdimension der Anspruchsvollen Aufgaben zu, welche während der Schulschließungen von den Schülerinnen und Schülern möglicherweise selbstbestimmter bearbeitet werden konnten und in differenzierter Form von der Lehrkraft bereitgestellt wurden. Zudem bestätigen die Ergebnisse, dass das Lernen in Zeiten des Distanzunterrichts weitaus stärker individuelle Unterschiede zwischen Schülerinnen und Schülern widerspiegelten.

A-14: Understanding and decreasing discrimination in school contexts

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* T03 R02 D81

Understanding and decreasing discrimination in school contexts

Chair(s): **Johanna Fee Ziemes** (University of Duisburg-Essen, Germany)
*Diskutant*in(nen):* **Kerstin Göbel** (University of Duisburg-Essen)

Discrimination can be described as the unfair treatment of a person due to their actual or pervieved membership (Jonas & Beelmann, 2009). Individuals face unfair treatment in supposedly safe spaces such as schools (Seaton et al., 2018). Previous research has shown that students' experiences of discrimination are negatively related to well-being and academic engagement (Huyhn, & Fulign, 2010; Verkuyten et al., 2019). In a meta-analysis of 218 studies on the negative consequences of discrimination (Benner et al. 2018), it was found that many measures of discrimination used for pupils are either old or developed for adults. Therefore, they might not adequately capture students' experiences of discrimination and their detrimental consequences. State of the art research needs tailored instruments to assess the levels of discrimination students experience, how these experiences relate to their psychosocial adjustment, and what teacher attitudes and behaviors can help to fight discrimination in school.

This symposium presents three approaches to improve understanding of discrimination in schools: The first presentation shows how an experience sampling design can be used to decrease retrospective bias when reporting interpersonal experiences of ethnic discrimination in the school context. By further differentiating between open and subtle forms of discrimination, this study investigated the differential effects discrimination experiences on behavioural, cognitive, and emotional school engagement. The second presentation demonstrates the value of open-ended questions in research on experiences of discrimination. Many measures focus on only one aspect of discrimination (e.g., ethnicity). The open question allowed students to identify multiple dimensions of discrimination that are rarely assessed and to assess how discrimination is connected to students' identification with the school. The final presentation focuses on the measurement of diversity related attitudes in (pre-service) teachers. A lack of positive attitudes towards cultural diversity and confidence to discuss culture-related issues can make those teachers more likely to discriminate against students or ignore issues of discrimination in the classroom. The authors show how an intervention implemented in initial teacher training can serve to improve the confidence of pre-service teachers to treat all students equally and to engage in culture-related conversations.

All three presentations offer new insights concerning current developments in measurement of and interventions against discrimination of students in school contexts. These findings have different applications and might therefore contribute to a differential understanding of marginalisation and its consequences. This understanding might provide avenues for interventions, teacher trainings, and structural changes that improve marginalized students' safety at school.

Beiträge des Symposiums

A daily diary study on associations between school-based ethnic discrimination and school engagement

Sauro Civitillo[1], Philipp Jugert[2], Tiffany Yip[3], Priscilla Lui[4], Peter Tietzmann[5], Haci-Halil Uslucan[2]
[1]Utrecht University, [2]University of Duisburg-Essen, [3]Fordham University, [4]Southern Methodist University, [5]Leibniz University Hanover

Being the target of discriminatory actions due to ethnicity can lead ethnic minority students to disengage from school (Verkuyten et al., 2019). Yet, research on links between school-based ethnic discrimination and school engagement has mostly examined between-person level associations (e.g., Leath et al., 2019). Typically, these studies relied on cross-sectional data with victims of discriminatory actions reporting their memories retrospectively, in the last year, or in a non-specific timeframe. This enhances the risk of retrospective bias in self-reported experiences (Goetz et al., 2013), leading respondents to over or under-estimate the occurrence of discrimination. Studies that have used daily diaries demonstrated that discrimination experiences vary from one day to the next, uncovering within-person fluctuations in the same ethnic group (e.g., English et al., 2020). Similar to discrimination, conceptualizing school engagement as a stable motivational force that only varies between students may not fully capture the phenomenon, as previous research has shown its context- and temporal-dependency, substantially varying at the within-student level (Patall et al., 2018).

We drew upon the conceptual model put forth by Verkuyten, Thijs, and Gharaei (2019), which highlights the interrelations between ethnic discrimination, social identity needs, and school engagement. This conceptual model posits that individuals who are exposed to discrimination may experience social identity threat. Discrimination can convey to individuals that they are devalued because of their ethnic group membership (Umaña-Taylor, 2016). Identity threat is problematic because group identities tend to satisfy a range of social identity needs (Breakwell, 1986). The identification with a particular group fulfills the needs of control (sense of efficacy) and belongingness (closeness to others). When the fulfillment of these identity needs is threatened, students may react by psychologically disengaging from school. In the present study, we aim to investigate daily-level within-person associations between discriminatory experiences and school engagement in a sample of students of immigrant descent (i.e., Turkish and Arab origin) across ten days.

Our sample included 87 students (62% female, M=15 years old) from secondary schools in the Ruhr area (44,7% Gesamtschule). Participants filled out 10 daily reports (from Monday to Friday) in the evening using a smartphone application over a 2-week period. Nine items measured school-based ethnic discrimination (three items for open discrimination by peers: e.g., ‘Today in school another student called me names or teased me because of my ethnic origin’; three items for subtle discrimination by peers: e.g., ‘Today in school another student did not interact with me in a conversation because of my ethnic origin’; and three items for teacher-based discrimination: e.g., ‘Today in school the teacher picked me less often because of my ethnic origin’). Twelve items measured daily behavioral, emotional, and cognitive school engagement.

On average, teacher-based discrimination experiences were significantly more frequent than open $t(956)=6.786$, $p<.001$, and subtle discrimination by peers $t(956)=6.774$, $p< 001$. There was no evidence for a significant difference between open and subtle discrimination by peers. In the multilevel models, after controlling for baseline measurements of discrimination and school engagement, gender, family heritage, ethnic self-identification, and family cultural capital, within-person effects indicate that on days students experienced teacher-based discrimination, they reported a decrease in cognitive ($\beta=-0.12$, $p<.05$) and in emotional school engagement ($\beta = -0.11$, $p<.05$). On days students experienced subtle discrimination by peers (but not open discrimination), they reported lower levels of

emotional engagement (β=-0.15, p<.01). Behavioral engagement was not predicted by any of the types of discrimination and sources. Findings from the current study inform interventions aimed at helping adolescents and educational practitioners to recognize and fight discrimination in schools.

Keywords: Ethnic discrimination, school engagement, secondary school, daily diary

Measureing students' experiences of discrimination in school

Johanna Fee Ziemes, Hermann Josef Abs
University of Duisburg-Essen

Thereotical Background

While many school teachers are motivated to decrease discrimination, the school system participates in the construction of differences that are the basis for discrimination (Pfaff & Tervooren, 2020). Discrimination can be understood as the perception of being illegitimately treated different because of ones (ascribed) group memberships (Jonas & Beelmann, 2009) and is associated with disadvantages in education (Hummrich, 2017; Scherr, 2017). The German law defines and bans six forms of discrimination (AGG, 2006): Discrimination based on ethnicity, gender, religion or ideology, disability, age, or sexual identity. Researchers stress that more forms of discrimination exists and name, for example, class based discrimination, discrimination based on beauty, and discrimination based on ones status as a child (IiP, 2015). Current measurement approaches tend to group specific modes of discrimination (e.g. Baumann, Egenberger & Supik, 2018), but for researchers and practitioners it is vital to know which aspects of discrimination adolescents experience and find relevant.

Research question

The authors investigate the practicality and validity of an open question on experiences of discrimination of adolescents. We will investigate what (ascribed) group memberships they connected with their experiences of discrimination in schools.

Methods

The study uses data from the German field trial of the International Civic and Citizenship Education Study 2022. Overall, 1,175 students from 23 schools participated. Students filled out a computer based questionnaire that employed a rotating booklet design. Therefore, not all students were presented with all questions. The questionnaire included a civic knowledge test and questions concerning students' attitudes, behaviors, and experiences at school

Students were presented with one definition of discrimination. On the next page, students were asked if they had experiences discrimination at their school within the last 12 months and if this discrimination was perpetrated by students, teachers, or both (single-choice). Students were also asked to write in an open text field what (ascribed) group membership connected to this discrimination.

Of the 779 students that saw this question 148 (19%) reported to have any experience with discrimination at school and 117 students (15%) gave an answer in the open text field. These answers were coded deductively along the categories of the AGG and the additional categories of the IiP. In a second step, deductive categories were created. Answers could be coded in multiple categories.

Results and their relevance

To validate the question on discrimination we investigated the relationship of students' experiences of discrimination and their feeling of belonging. A t-test revealed that students that experienced any form of discrimination were less likely to feel like they are part of their school (t(516)=-4.377; p=.004).

In the open text field students answered most frequently about discrimination based on their looks (N=30). Furthermore, 24 students reported being treated unfairly by their teachers; 18 reported experiences of racism; 9 from bad experience based on physical or psychological illnesses or disabilities; 8 wrote about discrimination based on their sexual orientation; 4 reported discrimination based on their religion or ideology, and 2 about poverty based discrimination.

The results show that a relevant number of students experience discrimination and that these experiences decrease school belonging. The categories of the German AGG were less relevant for students than other aspects of discrimination. Look-based discrimination was reported the most frequently. The runner up, (perceived) abuse of power by teachers, could be connected to adultistic attitudes. Further research is necessary here. The study as whole demonstrates the relevance of research concerning students' perception of discrimination. Open questions my help to capture experiences of discrimination that are relevant for students.

Keywords: discrimination, measurement approaches mixed method design, school identity

Training pre-service teachers for more inclusive and equitable classrooms – lessons learned from a project-based seminar

Sharleen Pevec[1], Linda Juang[1], Maja Schachner[2]
[1]University of Potsdam, [2]Martin-Luther-University Halle-Wittenberg

Although teachers are expected to be competent in culturally diverse classrooms (Standing Conference of the Ministers of Education and Cultural Affairs, 2013), they are not adequately prepared for this task (Morris-Lange, Wagner & Altinay, 2016; European Commission, 2017). This is particularly relevant as teachers play an important role when it comes to promoting equity and inclusion or fostering inequalities and exclusion in the classroom. The latter might show through ethnic discrimination in class (Götz, 2021), lower achievement expectations (Lorenz, 2021) and the risk of stereotype-threat for culturally minoritized students (Schofield & Alexander, 2012). Competent teachers, however, can be supportive, e.g. for students academic engagement (Bottiani et al., 2020), by providing opportunities for intergroup contact and reducing prejudice (Di Bernardo et al., 2017) and can be a protective factor for ethnically discriminated students (Civitillo et al., 2021).

To address the need of training pre-service teachers to promote equity and inclusion in the classroom, both relevant components of Culturally Responsive Teaching (Gay, 2015), we developed a project-based seminar for pre-service teachers in 2020. In this seminar, university students themselves complete and learn how to implement the Identity Project (Umaña-Taylor & Douglass, 2017; Juang et al., 2020) as a concrete way to become culturally responsive teachers. All project sessions are discussed, complemented with education and psychological theories, while encouraging university students' self-critical reflections on ethnic- and culture-related topics.

We hypothesized that in the intervention group (N = 266, Mage 22.64 years, 77% female), compared to the control group (N = 340, Mage 23.88 years, 82% female) would show more positive attitudes

towards culturally minoritized groups and inclusion as well as more self-efficacy for equitable teaching in culturally diverse classrooms over time. We measured pre-service teachers attitudes and self-efficacy through the scales of Deficit thinking about students (Fergus, 2016, 2017), Importance of including diversity in the curriculum (Dee & Henkin, 2002), Confidence in treating all students equally (Tropp & Rucinski, 2022) and Confidence engaging in ethnic- and culture-related conversations (Tropp & Rucinski, 2022).

Data in our quasi-experimental study were collected in week one (T1) and week ten (T2) in 11 intervention and 13 control seminars at two German universities. Findings of ANOVAs with repeated measures revealed significant time*group interactions for pre-service teachers confidence about treating all students equally ($F(1, 604) = 4.11$, $p < .043$, $\eta^2 = .007$) and engaging in ethnic- and culture-related conversations ($F(1, 604) = 6.67$, $p < .010$, $\eta^2 = .011$), increasing more in the intervention group from T1 to T2, but not for the measures of deficit thinking about students ($F(1, 604) = 2.49$, $p < .115$, $\eta^2 = .004$) and the importance of including diversity in the curriculum ($F(1, 604) = 0.47$, $p < .495$, $\eta^2 = .001$).

Our preliminary results indicate that pre-service teachers in both groups showed similar tendencies for more positive attitudes towards culturally minoritized groups and inclusion over time, but pre-service teachers in the Identity Project seminar gained more self-efficacy for equitable teaching in culturally diverse classrooms. Supporting future teachers to feel more confident in addressing ethnic- and culture-related issues, such as discrimination, as well as trying to make all students feel seen and valued, can be a first important step towards more equitable and inclusive classrooms. With repeated measures ANCOVAs, we will further test whether the findings remain stable.

Keywords: teacher training, cultural diversity, equity, culturally responsive teaching

A-15: Simulationsbasiertes Lernen und Messen in der Mediziner:innen und Lehrer:innenausbildung

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* T03 R03 D75

Simulationsbasiertes Lernen und Messen in der Mediziner:innen und Lehrer:innenausbildung

Chair(s): **Ann-Kathrin Schindler** (Universität Augsburg, Germany)
*Diskutant*in(nen):* **Ingo Kollar** (Universität Augsburg, Philosophisch-Sozialwissenschaftliche Fakultät, Lehrstuhl für Psychologie m.b.B.d. Pädagogischen Psychologie)

In der Ausbildung von Ärzt:innen und Lehrkräften sind simulationsbasierte Lernumgebungen ein adäquates Instrument, um Studierende schrittweise, ohne Handlungskonsequenz und effektiv an die Berufspraxis heranzuführen (Heitzmann et al., 2019). Gleichzeitig werden Simulationen genutzt, um Lern- und Denkprozesse der Nutzer:innen zu beforschen (Cook et al., 2013). Simulationen ermöglichen 1) Den Zugang zu im realen Feld schwer oder unökonomisch erfassbaren Forschungsgegenständen 2) die Überprüfung variierender instruktionaler Designs 3) eine Weiterentwicklung als Lernumgebung entlang gewonnener Erkenntnisse. In diesem Symposium stellen wir Beispiele simulationsbasierter Lern- und Messumgebungen aus der Ausbildungsforschung der Medizin und Lehrerbildung vor. Im Anschluss werden Synergien und Unterschiede der Disziplinen und Simulationsformen diskutiert.

1. N=150 Lehramtsstudierende schätzten in einer *video-basierten Simulation* zur Messung von Diagnosekompetenzen Schüler:innen ein. Entlang zusätzlich erhobenen Wissens, Interesse, Selbstregulation und Selbstwirksamkeit wurden Lernvoraussetzungsprofile der Studierenden abgeleitet. Es konnten u.a. tentative Unterschiede für drei Lernvoraussetzungsprofile der Studierenden (wissend; motiviert; potenziell gefährdet) in der Diagnoseakkuratheit gezeigt werden. Die Ergebnisse werden genutzt, um die Simulation mit Prompts anzureichern, um Studierende mit ungünstigeren Voraussetzungen im Erwerb von Diagnosekompetenz zu unterstützen.

2. N=57 Lehramtsstudierende durchliefen eine *virtual reality (VR) Simulation* für den Unterrichtseinsatz mit Schüler:innen zum exponentiellen Wachsum von Infektionskrankheiten. Zusätzlich wurde das Potenzial und die wahrgenommene Relevanz von VR Anwendungen für den Unterricht pre-post erfragt. Ein Anstieg der wahrgenommenen Nützlichkeit von VR gibt Hinweise, dass das aktive Durchlaufen der Simulation eine Anwendung in der späteren Unterrichtspraxis positiv beeinflussen kann.

3. N=67 Medizinstudierende im praktischen Jahr wurden in einer *realitätsnahen Arbeitsplatz-basierten Simulation*, die den ersten Tag der Facharztweiterbildung widerspiegelt, von Supervisoren bezüglich ausgewählter Kompetenzfacetten eingeschätzt. Anschließend gaben die Supervisoren an, auf welchem Supervisionslevel sie den Studierenden bestimmte professionelle Tätigkeiten anvertrauen würden. Die Ergebnisse zeigen, dass der Beobachtungsfokus auf konkrete Kompetenzfacetten eine Grundlage für Anvertrauensentscheidungen darstellen kann. In einer zweiten Teilstudie wurden von N=87 Studierenden die Patientenübergaben in der Simulation videographiert und sprachlich analysiert. Informationsstärkende und -abschwächende Ausdrücke bezüglich Diagnostik- und Therapieentscheidungen kamen am häufigsten bei herausfordernden Patientenfällen vor.

4. In einer *avatar-basierten Simulationsumgebung* schätzten N=26 Medizinstudierende ein, welche professionellen Tätigkeiten sie durch Avatare verkörperten Personen anvertrauen würden. Die

Avatare zeigten dabei drei Gesichtsausdrücke modelliert entlang des Facial Action Coding System nach Ekman und Friesen (2002) (neutral, authentisch lächeln, nicht-authentisch lächeln). Authentisch lächelnden Avataren wurden Tätigkeiten signifikant häufiger anvertraut als welchen mit neutralem Ausdruck oder unauthentischem Lächeln. Die Ergebnisse zeigen, dass die Art des Lächelns einen Einfluss auf Anvertrauensentscheidungen zu haben scheint und dienen als Grundlage für die Weiterentwicklung der Simulationsumgebung für Ausbildner:innen variierender Expertise.

Beiträge des Symposiums

Erfassung von Diagnosekompetenzen in Simulationen: Einfluss der Charakteristika des Diagnostizierenden

Michael Nickl[1], Daniel Sommerhoff[2], Elias Codreanu[1], Stefan Ufer[3], Tina Seidel[1]
[1]Technische Universität München (TUM), TUM School of Social Sciences and Technology, Lehrstuhl für Pädagogische Psychologie, [2]IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel, [3]Ludwig-Maximilians-Universität (LMU), Mathematisches Institut der Universität München

Keywords

Diagnosekompetenzen, video-basierte Simulation, Lehrkräfteausbildung

Theoretischer Hintergrund

Täglich müssen Lehrkräfte den Kenntnisstand ihrer Schüler:innen diagnostizieren. Gute Diagnosekompetenzen, beispielsweise im Sinne von akkuraten Diagnosen, sind daher wichtiges Merkmal einer „guten“ Lehrkraft (KMK, 2004). Allerdings ist die Diagnoseakkuratheit von Lehrkräften teils verbesserungswürdig (Urhahne & Wijnia, 2021). Daher erscheint es wichtig, Diagnosekompetenzen bereits im Studium zu fördern (Heitzmann et al., 2019). Zur Messung von Diagnosekompetenzen, die einer bedarfsorientierten Förderung vorangeht, sowie zur Förderung selbst haben sich Simulationen als effizient und effektiv gezeigt (Chernikova et al., 2020).

Die Messung von Diagnosekompetenzen in der bisherigen Forschung, oft mithilfe textbasierter Tests, offenbarte einen Einfluss der professionellen Wissensbasis und motivational-affektiver Charakteristika der Diagnostizierenden auf die Diagnoseakkuratheit (Ostermann et al., 2018; Westphal et al., 2018). Dies wird auch für Simulationen angenommen und in der vorliegenden Studie empirisch untersucht (Heitzmann et al., 2019).

Fragestellungen

Welche Rolle haben individuelle Lernvoraussetzungen von angehenden Lehrkräften in einer Simulation zur Messung von Diagnosekompetenzen? Genauer: Lassen sich unterschiedliche Profile kognitiver und motivational-affektiver Charakteristika von Lehramtsstudierenden in der Simulation ermitteln? Falls ja: Welche Profile lassen sich identifizieren? Unterscheiden sich diese Profile in Diagnoseprozess und/ oder Diagnoseakkuratheit?

Methode

Zur Untersuchung der Fragestellungen abolvierten N = 150 Lehramtsstudierende (67% weiblich, Semester: M=5,0; SD=2,5), eine ca. 90-minütige Simulation. Zur Erfassung der individuellen Lernvoraussetzungen wurden Skalen zu professioneller Wissensbasis, individuellem Interesse, Selbstregulation und Selbstwirksamkeit erhoben. Die Erfassung der Diagnoseakkuratheit erfolgte durch Likert Ratings innerhalb der video-basierten Simulation, wobei die Studierenden mit jeweils acht

Items zwei Schüler:innen der 7. Klasse einschätzten. Der Diagnoseprozess, in dem die Studierenden bis zu zehn Videos mit den beiden Schüler:innen sehen konnten, wurde auf Basis schriftlicher Notizen untersucht, die die Studierenden während der Beobachtung der Videos machten. Alle Skalen wurden bereits in früheren Studien eingesetzt (Codreanu et al., 2020).

Die so gewonnenen Daten wurden mithilfe einer latenten Profilanalyse (LPA) und der BCH Methode analysiert (Bolck et al., 2004; Muthén & Muthén, 1998-2017).

Ergebnisse und Diskussion

In der LPA deuteten die Fitindizes drei unterschiedliche Profile der Lernvoraussetzungen an: Ein Profil mit guter professioneller Wissensbasis und durchschnittlichen motivational-affektiven Charakteristika („wissend"), ein „motiviertes" Profil mit durchschnittlichem Vorwissen, aber hohen motivational-affektiven Charakteristika, sowie ein „potentiell gefährdetes" Profil, das jeweils unter dem Durchschnitt lag.

Hinsichtlich der Diagnose zeigte sich jeweils dieselbe deskriptive Rangordnung: Das wissende Profil wies die höchsten und das potentiell gefährdete Profil die geringsten Werte auf. Inferenzstatistisch zeigten sich signifikante Unterschiede in der Anzahl der gesehenen Videos ($\chi2 = 7,12$; $p = 0,028$), aber nicht hinsichtlich der Anzahl der geschriebenen Wörter in den Notizen ($\chi2 = 3,37$; $p = 0,124$) und der Diagnoseakkuratheit ($\chi2 = 4,30$; $p = 0,116$). Die Vergleiche von wissendem und potentiell gefährdetem Profil waren hinsichtlich Anzahl der Videos ($\chi2 = 6,06$; $p = 0,014$) und Diagnoseakkuratheit ($\chi2 = 4,13$; $p = 0,042$) waren dabei signifikant.

Die tentativen Unterschiede der drei Lernvoraussetzungsprofile in der Diagnoseakkuratheit sowie deren Unterschiede im Diagnoseprozess unterstreichen die Rolle der Lernvoraussetzungen beim Diagnostizieren (Heitzmann et al., 2019): Während der Vorteil des wissenden Profils gegenüber dem potentiell gefährdeten Profil auf den positiven Einfluss einer guten professionellen Wissensbasis hindeutet (Ostermann et al., 2018), scheinen die Einflüsse hoher motivationaler Lernvoraussetzungen nur tentativ wie der Vergleich des motivierten und des potentiell gefährdeten Profils zeigt. Dies könnte daran liegen, dass das motivierte Profil dem Diagnoseprozess im Vergleich zum wissenden Profil zu wenig Aufmerksamkeit schenkt (Herppich et al., 2018). Versteht man die Simulationsumgebung als Lernumgebung, in die sich instruktionale Unterstützung wie Scaffolding einbetten lässt (Chernikova et al., 2020). Die identifizierten Lernvoraussetzungsprofile bieten hier einen Anhaltspunkt für adaptive Unterstützung, beispielsweise in Form von Prompts, die das motivierte oder das potentiell gefährdete Profil im Diagnoseprozess unterstützen. Die Untersuchung solch adaptiver Unterstützung bedarf weiterer Forschung (Plass & Pawar, 2020).

Wie nehmen Lehramtsstudierende eine neuartige VR-Umgebung für Schüler:innen wahr? Eine Mixed-Methods-Studie zur Veränderung von Nützlichkeitsüberzeugungen

Florentine Hickethier, Mathias Dehne, Alexander Gröschner
Friedrich-Schiller-Universität Jena, Fakultät für Sozial- und Verhaltenswissenschaften, Institut für Erziehungswissenschaft, Lehrstuhl für Schulpädagogik und Unterrichtsforschung

Keywords

Simulationsbasiertes Lernen, Lehrerbildung, Virtual Reality

Theoretischer Hintergrund

Koppenberg et al. (2014) beschreiben simulationsbasiertes Lernen (SBL) als eine erfahrungsorientierte Form des Wissenszuwachses, durch die konkrete Aufgaben unterschiedlich realistisch dargeboten und von Lernenden erprobt werden. Diese Form des Lernens bietet auch für die Lehrerbildung großes Potenzial und erste Arbeiten liefern Belege für deren Wirksamkeit in Bezug auf den Kompetenzzuwachs von Studierenden des Lehramts (Chernikova et al., 2020).

Der Theory of Reasoned Action (TRA; Davis, 1989) zufolge werden die Benutzerfreundlichkeit und die wahrgenommene Nützlichkeit als zentrale Einflussfaktoren angesehen, die das Nutzungsverhalten gegenüber neuen Technologien bedingen. Aus dieser Theorie ergibt sich die Frage, inwiefern sich die Nützlichkeitsüberzeugungen verändern, wenn Lerner:innen an Erfahrung im Umgang mit der entsprechenden Technologie zugewinnen. Die aktuelle Studie erprobt den Einsatz von Virtual Reality (VR) in der Lehrerbildung.

Fragestellungen

Im Rahmen eines EFG-Projekts (Azevedo et al., 2020) gehen wir u.a. der Frage nach, wie angehende Lehrpersonen eine VR-Lernumgebung wahrnehmen, die für den mathematisch-naturwissenschaftlichen Unterricht bestimmt ist und in dem eine mathematische Aufgabe zum exponentiellen Wachstum simulationsbasiert gelöst werden soll. In der vorliegenden Teilstudie untersuchen wir die Einschätzungen der Studierenden zur Nützlichkeit und Anwendbarkeit des VR-Tools in der Unterrichtspraxis und welche Emotionen im Rahmen der VR-Erfahrung Studierende beschreiben.

Methode

Zur Umsetzung der EFG-Studie wurde im Sommersemester 2021 eine VR-Aufgabe für Schüler:innen simulationsbasiert mit Lehramtsstudierenden erprobt (Pandemic by prisms). Die VR-Umgebung thematisiert die Ausbreitung von Infektionskrankheiten. Über die Erforschung von Eindämmungsstrategien soll Schüler:innen exponentielles Wachstum nahegebracht werden. Das Durchlaufen der VR-Umgebung nahm im Durchschnitt etwa 45 Minuten in Anspruch. Im Voraus und im Anschluss an die VR-Aufgabe bearbeiteten die Teilnehmer:innen einen Fragebogen unter anderem mit den Skalen zur Perceived Usefulness (Teo, 2009), zu den Success Expectations (Rheinberg et al., 2001), zur Situational Perception of the Task (Wigfield, 1994) sowie Epistemic Emotions (Pekrun et al., 2016). Die Skala von Teo (2009) zur Perceived Usefulness von VR wurde im Posttest noch durch die Skalen Ease of Use sowie Intention to Use erweitert. Alle Items wurden auf einer fünfstufigen Likert-Skala eingeschätzt (1 = „trifft überhaupt nicht zu“ bis 5 = „trifft voll und ganz zu“). Zusätzlich wurden am Ende eines jeden Durchlaufs mit den Studierenden Einzelinterviews durchgeführt, in denen ebenfalls das Potenzial und die wahrgenommene Relevanz simulationsbasierter Anwendungen für den Unterricht erfragt wurden.

An der vorliegenden Studie nahmen insgesamt 57 angehende Lehrpersonen aus Finnland, Deutschland, Israel und der Türkei teil (Alter: M = 26.18 Jahre, SD = 4.65 Jahre; Geschlecht: 36.8% weiblich, 57.9% männlich, 5.3% ohne Angabe). 38.6 % der Teilnehmenden hatten zuvor noch nie eine VR-Umgebung ausprobiert.

Ergebnisse und Diskussion

Hinsichtlich der quantitativen Befunde zeigen erste Ergebnisse kleine bis mittlere Effekte in den affektiv-motivationalen Variablen von Prä- zu Posttest in Bezug auf die gesamte Stichprobe (z.B. wahrgenommener Nutzen, Beispielitem: „Der Einsatz von VR wird meine Effektivität erhöhen“; z = -2,14, p = 0,032, r = 0,2). Ähnliches gilt für die positiven Emotionen (z.B. Neugier) der Lehramtsstudierenden während der VR-Erfahrung, für die ebenfalls statistisch signifikante Prä-Post-Unterschiede auftraten (z = 4,22, p < 0,001, r = 0,3). Die Erkenntnisse über die affektiv-motivationalen Variablen und die empfundenen Emotionen stützen folglich die Aussagen des Technology Acceptance Modells (TAM; Davis et al., 1989). Ein Anstieg der wahrgenommenen Nützlichkeit von VR, bedingt durch den Umgang mit dieser Technologie, beeinflusst die Einstellungen gegenüber der Anwendung in der späteren Unterrichtspraxis positiv.

Hinsichtlich der Kodierung der Interviews zeigen sich gute Interrater-Reliabilität aller Kodierungen (.70 < Cohen’s κ < .80). Die quantitativ sowie qualitativ gewonnenen Ergebnisse werden im Kontext der Diskussion um Wirksamkeit VR-basierter Simulationsumgebungen in der Lehrerbildung diskutiert. Für die zukünftige Forschung soll der Frage nachgegangen werden, inwiefern simulationsbasierte Settings in der Lehramtsausbildung Anwendung finden können, um Lehrpersonen im Universitätskontext auf die Unterrichtspraxis an den Schulen vorzubereiten.

Anvertrauen von Verantwortung auf Basis von Arbeitsplatz-basierten Beobachtungen

Sigrid Harendza[1], Martina Kadmon[2], Pascal Berberat[3], Fabian Fincke[4], Julia Gärtner[1], Kristina Schick[3]
[1]Zentrum für Innere Medizin, Sektion Ausbildungsforschung, Universitätsklinikum Hamburg-Eppendorf, [2]Universität Augsburg, Department of Medical Education Augsburg (DEMEDA), [3]Technische Universität München (TUM), TUM Medical Education Center, Lehrstuhl für Medizindidaktik, medizinische Lehrentwicklung und Bildungsforschung, [4]Evangelisches Krankenhaus Oldenburg, Universitätsmedizin Oldenburg

Keywords

Anvertraubare Professionelle Tätigkeiten (APT), Anvertrauensentscheidungen, Arbeitsplatz-basierte Simulation, sprachliche Expression von Unsicherheit

Theoretischer Hintergrund

Anvertraubare professionelle Tätigkeiten (APT) stellen eine kompetenzorientierte Arbeitsplatz-basierte Trainings- und Assessment-Rahmenstruktur dar, welche die schrittweise Verantwortungsübertragung an Lernende und Reduktion der Supervision durch Lehrende bei der Durchführung beruflicher Tätigkeiten beinhaltet (Meyer et al., 2019). Der Supervisionsgrad reicht von direkter Beobachtung über bedarfsorientierte Verfügbarkeit des Lehrenden bis zur umfassenden Verantwortungsübergabe (ten Cate et al., 2015). Eine APT umfasst verschiedene Kompetenzfacetten (FOC = facets of competence), die eine fokussierte Beobachtung ermöglichen (Berberat et. al., 2013; ten Cate et al., 2015). Das APT-Konzept beinhaltet neben ad hoc-Anvertrauensentscheidungen in konkreten Beobachtungssituationen (ten Cate et al., 2015) auch prospektive Entscheidungen (Fincke et al., 2020). Lehrende in der Medizin treffen Anvertrauensentscheidungen für Tätigkeiten an Patient:innen, bei denen Subjektivität und ein Maß an Unsicherheit unvermeidbar sind (Rotthoff et al., 2021). Kriterien für fundierte Anvertrauensntscheidungen müssen erforscht werden, um die sichere Umsetzung am Arbeitsplatz zu gewährleisten.

Fragestellungen

Der Beitrag orientiert sich an zwei zentralen Fragestellungen:

1. Wie beeinflusst die Bewertung konkreter Kompetenzfacetten (FOC-Score) Lehrende in ihren prospektiven Anvertrauensentscheidungen von APT (APT-Score)? (Studie 1)

2. Sind implizite sprachliche Modifikatoren, z.B. in simulierten Übergabesituationen von Patient:innen, geeignet, um Lehrende in ihren Anvertrauensentscheidungen zu unterstützen? (Studie 2)

Methode

Studie 1: Studierende im Praktischen Jahr des Medizinstudiums oder kurz davor (n=67) nahmen an einem simulierten, realitätsnahen Arbeitsplatz-basierten Assessment teil, das den ersten Tag der Facharztweiterbildung widerspiegelt (Prediger et al., 2020). Die Simulation bestand aus drei Phasen: einer Konsultation durch 5 Schauspielpatient:innen, einer individuellen Phase zur Erarbeitung eines Managementplans und Patientenübergaben an reale Ärzt:innen, die videografiert und transkribiert wurden. Neun erfahrene Kliniker:innen standen den Studierenden als Supervisor:innen zur Verfügung, trafen sich in jeder Phase einmal mit ihrem/r zugeteilten Studierenden und beobachteten die simulierte Patientenübergabe. Auf Basis der gesamten Simulation bewerteten sie ihre/n Studierenden bezüglich sieben Kompetenzfacetten. Darüber hinaus trafen sie prospektive Anvertrauensentscheidungen für 12 APT, die keinen direkten Bezug zur Simulation hatten, aber jeweils mehrere Kompetenzfacetten beinhalteten (Prediger et al., 2019; Prediger et al., 2020). Analysiert wurde der Zusammenhang zwischen FOC-Scores und APT-Scores durch Analyse der Pearson-Korrelationskoeffizienten sowie der Einfluss einzelner FOC-Scores auf die Anvertrauensentscheidungen über alle APT (Gesamt-APT-Score) durch multiple lineare Regressionsanalyen. Die statistischen Analysen erfolgten durch R 3.4.3 (Fincke et al., 2020).

Studie 2: Inkludiert wurden 87 Studierende, die das simulierte Assessmentformat mit jeweils 3 Schauspielpatient:innen durchliefen und deren Patientenübergaben videografiert und transkribiert wurden. Alle verbalen Sequenzen, die im Rahmen eines zuvor etablierten Bezugssystems von impliziten Ausdrücken verortet waren, und medizinische Informationen abschwächten oder verstärkten und Unsicherheit bzw. Sicherheit suggerierten (Gärtner et al., 2020), wurden durch eine Wortsuche mittels MAXQDA identifiziert.

Ergebnisse und Diskussion

Fragestellung 1: FOC- und APT-Scores zeigten eine signifikante positive Korrelation über alle FOC und APT hinweg (Pearson $r=0,72$, $p<0,001$) (Fincke et al., 2020). Geringere signifikante positive Korrelationen wurden zwischen einzelnen FOC-Scores und dem Gesamt-APT-Score beobachtet ($p<0,01$), während individuelle APT-Scores nicht mit allen einzelnen FOC-Scores korrelierten. Diese Ergebnisse stützen die Annahme, dass Anvertrauensentscheidungen verschiedene Kompetenzfacetten widerspiegeln und der Beobachtungsfokus auf Kompetenzfacetten eine Grundlage für Anvertrauensentscheidungen darstellen kann. Die Regressionsanalyse zeigte, dass zwei Kompetenzfacetten „Teamarbeit und Kollegialität“ (FOCa) und „Struktur, Arbeitsplanung und Priorisierung“ (FOCb) den Gesamt-APT-Score am besten vorhersagten (lineare Prädiktion APT-Score = 0,46 FOCa + 0,42 FOCb).

Fragestellung 2: Die sprachliche Analyse der videografierten Patientenübergaben ergab die Nutzung von insgesamt 3384 verstärkenden bzw. abschwächenden Modifikatoren. In Übergabesituationen von

besonders komplexen Patientenfällen wurden informationsstärkende und -abschwächende Ausdrücke bezüglich Diagnostik- und Therapieentscheidungen am häufigsten verwendet, z.B. 757 mal (22,4%) bei einem aggressiven Patienten mit starken Bauchschmerzen (Gärtner et al., 2022). Es soll diskutiert werden, ob implizite sprachliche Hinweise auf Unsicherheit in der klinischen Argumentation von Handlungs- und Entscheidungsprozessen Einfluss auf Anvertrauensentscheidungen von Lehrenden nehmen können.

Einfluss von Mimik in medizinischen „ad-hoc"-Anvertrauensentscheidungen – Pilotstudie zu einer Avatar-basierten Simulationsumgebung

Moritz Bauermann[1], Ann-Kathrin Schindler[1], Miriam Kunz[2], Elisabeth André[3], Thomas Rotthoff[1]
[1]Universität Augsburg, Department of Medical Education Augsburg (DEMEDA), Lehrstuhl für Medizindidaktik und Ausbildungsforschung, [2]Universität Augsburg, Department of Medical Education Augsburg (DEMEDA), Lehrstuhl für medizinische Psychologie und Soziologie, [3]Universität Augsburg, Fakultät für Angewandte Informatik, Lehrstuhl für Menschzentrierte Künstliche Intelligenz

Keywords

Anvertraubare Professionelle Tätigkeiten (APT), Anvertrauensentscheidungen, Avatar-basierte Simulationsumgebung, Mimik

Theoretischer Hintergrund

Im Rahmen des Studiums werden angehende Mediziner:innen an klinische Tätigkeiten herangeführt. Dem Konzept der Anvertraubaren Professionellen Tätigkeiten (APTs) kommt hier große Bedeutung zu. APTs sind klinische Arbeitseinheiten, welche Mediziner:innen in Abhängigkeit des Fortschritts ihrer Ausbildung zur selbstständigen Durchführung anvertraut werden können (ten Cate, 2005). Ob eine ATP von einem Auszubildenden ausgeführt wird, hängt situationsspezifisch von der individuellen Entscheidung der ausbildenden Person (Trustor) ab (ten Cate et al., 2020). Der Trustor durchläuft einen Abwägungsprozess entlang aller vorliegenden Informationen, um maximale Patientensicherheit und -versorgungsqualität zu gewährleisten (ten Cate & Schumacher, 2022). In der klinischen Realität erfolgen Anvertrauensentscheidung häufig just im Moment für den möglicherweise unbekannten Auszubildenden (ten Cate et al., 2016). Bisher ist jedoch noch kaum erforscht, wie diese sogenannten „ad-hoc"-Anvertrauensentscheidungen zustande kommen. EnTRUST ist die Pilotstudie mit Medizinstudierenden für eine nachfolgende Hauptstudie mit Ärzt:innen. Im Rahmen der Pilotstudie wurde eine Simulationsumgebung entwickelt, in welcher Medizinstudierende professionelle Tätigkeiten an durch Avatare simulierte (Mit-)Studierende anvertrauen.

Fragestellung

Wir postulieren, dass die Gesichtsmimik eine wichtige Basis dafür ist, ob Personen ad-hoc (an)vertraut wird (Krumhuber et al., 2007). In der Pilotstudie untersuchen wir folgende Hypothese: Bei durch Avatare simulierten Personen mit authentischem Lächeln werden mehr positive Anvertrauensentscheidungen getroffen als bei nicht authentischem Lächeln oder neutralem Ausdruck. Gesichter, bei denen die Mundwinkel nach oben gezogen und deren Augen zusammengekniffen sind, gelten als authentisch lächelnd. Werden jedoch nur die Mundwinkel nach oben gezogen, wird das Lächeln als weniger authentisch erlebt (Fernández-Martín et al., 2017).

Methode

Entwicklung der Simulationsumgebung

Für die Simulationsumgebung wurden neun unterschiedliche, in ihrer Komplexität vergleichbare klinische Szenarien zutreffend für die Stichprobe der Pilotstudie formuliert. Beispielsweise kann ein Blutzuckertest aufgrund einer Terminüberschneidung nicht sofort selbst ausgeführt werden und muss einem Mitstudierenden anvertraut werden. Für die Erstellung der Avatare wurden Fotos von natürlichen Personen mittels der Software FaceGen Modeller eingescannt und in Avatare transformiert. Um den Einfluss individueller morphologischer Merkmale der Avatare auf die „ad-hoc"-Anvertrauensentscheidung zu reduzieren, wurden unterschiedliche phänotypisch kaukasisch aussehende Avatare generiert, die sich in Geschlecht, Aussehen, Frisur und Haarfarbe unterscheiden, bzgl. der Mimik jedoch standardisiert gestaltet wurden. Für die Modulierung des Lächelns wurde das Facial Action Coding System nach Ekman und Friesen angewendet (Ekman et al., 2002). Pro Ausdruck (authentisches Lächeln, nicht authentisches Lächeln, neutraler Gesichtsausdruck) wurden drei Avatare jeweils in weiblicher und männlicher Form erstellt (3 mimische Ausprägungen x 3 Avatare x 2 Geschlechtsausprägungen =18 Avatare). Zuletzt wurden die Avatare mithilfe der Software FACSGen als Video animiert (jeweils 10 Sek.).

Design und Stichprobe

26 Medizinstudierenden (65,4% weiblich; Alter: n20 – 25y = 25; n>26y = 1; Semester: n2tes = 9, n4tes = 16, nfehlend = 1) wurden neun Anvertrauensszenarien zweimal präsentiert, um eine Anvertrauensentscheidung für die 18 unterschiedlichen Avatare abzufragen. Am Ende jedes Szenarios trafen die Teilnehmer:innen eine Anvertrauensentscheidung (ja/nein). Zum Abschluss der 40ig minütigen Simulation wurde der demografische Hintergrund erfasst.

Auf Unterschiede im Anvertrauen zu wurden mit Varianzanalysen nach Friedman in SPSS Version 28 getestet.

Ergebnisse und Diskussion

Die ersten Auswertung ergab einen signifikanten Unterschied für die getroffenen Anvertrauensentscheidungen in Abhängigkeit des Gesichtsausdrucks (Friedman-Test: Chi-Quadrat(2) = 27,48, $p < 0,001$; Ränge: neutral = 1,56; nicht authentisches Lächeln = 1,67; authentisches Lächeln = 2,77). Post-hoc (Dunn-Bonferroni-Test) zeigte sich, dass authentisch lächelnden Avataren Tätigkeiten signifikant häufiger ad-hoc anvertraut wurden als welchen mit neutralem Ausdruck ($z = -1,21$, $p < 0.001$, $r = 0,24$) oder unauthentischem Lächeln ($z = -1,10$, $p < 0,001$, $r = 0,22$).

Die Ergebnisse mit den Medizinstudierenden zeigen, dass die Art des Lächelns einen Einfluss auf Anvertrauensentscheidungen zu haben scheint. Dies soll in der Weiterentwicklung mit ausbildenden Ärzt:innen getestet werden. Dafür wird die Simulationsumgebung hinsichtlich ihrer Szenarien weiter entwickelt und mit Proband:innen unterschiedlicher Ausbildungsexpertise getestet.

A-16: Ursachen und Wirkungen irregulärer Bildungsverläufe: Empirische Befunde aus Deutschland und Luxembourg

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* T03 R03 D89

Ursachen und Wirkungen irregulärer Bildungsverläufe: Empirische Befunde aus Deutschland und Luxembourg

Chair(s): **Paul Fabian** (Technische Universität Dortmund, Germany), **Alyssa Laureen Grecu** (Technische Universität Dortmund, Germany)
*Diskutant*in(nen):* **Steffen Schindler** (Universität Bamberg)

Der Eintritt in die institutionalisierte frühkindliche Bildung bildet in den meisten Fällen den Start einer langen Reihe von Übergängen zwischen verschiedenen formalisierten Bildungskontexten. Die soziale Selektivität im deutschen Schulsystem ist hoch (Ditton & Maaz 2011), was ebenfalls mit unterschiedlichen Bildungsverläufen (z.B. spätere Schuleintritte Klassenwiederholungen oder Schulwechsel) zwischen Schüler*innen verschiedener sozialer Herkunft verbunden ist. Ein irregulärer Bildungsverlauf kann als kritisches Lebensereignis (Fabian, 2020) verstanden werden und wirkt sich auf die Leistung von Schüler*innen, ihre sozioemotionale Entwicklung sowie langfristig auch auf Möglichkeiten zur Berufsausbildung und die gesellschaftliche Teilhabe aus (Ehmke et al. 2017; Marsh et al. 2017).

Bildungsverläufe hängen unter anderem von den Leistungen, der Motivation und dem Verhalten der Schüler*innen ab (Goos, et al., 2013). Neben diesen individuellen Faktoren prägen jedoch auch institutionelle und soziale Faktoren, wie beispielsweise die Durchlässigkeit des Bildungssystems oder die Förderung und Erwartungen einer Lehrperson, die Bildungsverläufe der Schüler*innen (Goos et al., 2013; König & Darge, 2010; Xia & Kirby, 2009). Nicht zuletzt wirken sich auch elterliche Entscheidungen über Verläufe und Wechsel im Bildungsverlauf ihrer Kinder aus, wobei die Entscheidungen auf Basis rationaler Abwägungen und der vorhandenen Ressourcen (Boudon 1974, Bourdieu 1982, Breen & Goldthorpe 1997, Esser 1999) getroffen werden.

Aufgrund der weitreichenden Konsequenzen von Bildungsverläufen und Übergangsentscheidungen ist es wichtig diese Negativspirale zu durchbrechen. Das vorliegende Symposium trägt hierzu bei, indem irregulärer Bildungsverläufe und deren Ursachen und Konsequenzen auf Basis internationaler erforscht werden und somit zum Verständnis von Bildungsverläufen beiträgt.

Der erste Beitrag widmet sich der sozialen Segregation unter den Bedingungen der freien Schulwahl in der Schuleingangsphase. Die Datenbasis für die Untersuchung einer mittelgroßen deutschen Stadt stellen Angaben zu Wohnbezirken, Schulen sowie Merkmalen des sozioökonomischen Hintergrundes und der Migrationsgeschichte der Schüler*innen dar und umfassen den kompletten Einschulungsjahrgang des Jahres 2021 (*N=3.243*).

Am Beispiel Luxemburgs erforscht **der zweite Beitrag** zunächst Verbindungen von Klassenwiederholungen mit weiteren Bildungsereignissen auf Basis eines administrativen Panel-Schülerdatensatz (Geburtsjahrgang 1990, Vollerhebung mit *N*= 3.375). In diesem Zuge werden Gruppen von Schüler*innen identifiziert, die besonders häufig von irregulären Bildungsverläufen betroffen sind. Darauf aufbauend liefern die Ergebnisse der narrativen Interviewstudie mit (ehemaligen) Schüler*innen, die selbst irregulär Bildungsverläufe erlebt haben, Einblicke darin wie Klassenwiederholungen erlebt werden.

Der dritte Beitrag fokussiert die Effektivität von Klassenwiederholungen, indem erforscht wird inwiefern eine Klassenwiederholung die Wahrscheinlichkeit erhöht ein schulformadäquates Zertifikat zu erreichen. Dazu werden Daten von Sekundarschüler*innen des deutschen nationalen Bildungspanels (NEPS; SC3, *N*=4.736) analysiert.

Zum Schluss werden die Beiträge diskutiert und eingeordnet.

Beiträge des Symposiums

Soziale Segregation beim Eintritt in die Grundschule – wo beginnen irreguläre Bildungsverläufe?

Johanna Gold, Annette Textor, Jan Christoph Störtländer
Universität Bielefeld

Die Übergänge zwischen Bildungsinstitutionen markieren zentrale Gelenkstellen in Bildungsbiografien. Der Umstand, dass durch Segregationsmechanismen an diesen Übergängen soziale Disparitäten verstärkt oder gar hervorgebracht werden, ist im bildungswissenschaftlichen Diskurs – insbesondere für den Übergang zwischen Primarstufe und Sekundarstufe I – breit erforscht(Z. B. Maaz et al. 2010). Kinder aus sozio-ökonomisch schwächer gestellten Haushalten sind im Hinblick auf die Anforderungen der Schule häufig stärker benachteiligt als Kinder aus sozio-ökonomisch stärkeren Familien und sind somit stärker der Gefahr irregulärer Bildungsverläufe – z-B. Klassenwiederholungen – ausgesetzt.

Segregation zeigt sich aber auch schon beim ‚ersten Übergang', dem Eintritt in die Grundschule. Sowohl bei verbindlichen Schulbezirken wie auch bei der Möglichkeit der freien Wahl der Grundschule ist eine Segregation der Schüler*innen nach sozialer Herkunft zu beobachten (Krüger, Breidenstein & Roch 2020). Diese Art der Segregation ist insbesondere aus zwei Gründen als problematisch zu betrachten.

Zum einen hat die Komposition von Schulen und Klassen einen Einfluss auf die Schullaufbahn der Kinder(Dumont et al. 2013). Die Schüler*innen bringen verschiedene Startvoraussetzungen, Erfahrungen und Merkmale mit. Als wichtigste Merkmale werden der Sozialstatus der Familien, familiäre Belastung, kulturelle Ressourcen, ethnisch-kulturelle Merkmale, lernbiographische Belastungen und kognitive Grundfähigkeiten angesehen (Baumert et al. 2006). Kommen in einer Klasse beson-ders viele Kinder mit schwierigen Startvoraussetzungen zusammen, können diese Kompositionen Einflüsse z.B. auf die Gestaltung des Unterrichts, die Interaktion der Schüler*innen oder die materielle Ausstattung der Schule nehmen. Dies wiederum kann zu einer verstärkten Benachteiligung von Kindern mit ohnehin schwierigen Startvoraussetzungen führen. Insbesondere bei Schü-ler*innen aus belasteten Lebenslagen können sich so zusätzlich zu den schon vorhandenen Be-nachteiligungen weitere schwerwiegende Nachteile ergeben.

Zum anderen engt die Segregation nach sozialer Herkunft die Möglichkeit der Schule ein, die Schüler*innen zu einer freien und reflektierten Wahl zwischen unterschiedlichen Typen der Lebensfüh-rung zu befähigen. Attributionen von Verwirklichungschancen und Vorstellungen von individuell zu erreichenden Bildungsabschlüssen sind in einem hohen Maß abhängig von der habituellen Prägung der Familie (Bourdieu 1982). Ein entscheidendes Maß für die geglückte Verwirklichung eines Lebensentwurfes ist nach dem Capability Approach die Menge und die Qualität des Spektrums von-

einander unterscheidbarer Möglichkeiten und Fähigkeiten von Menschen (Otto und Ziegler 2008). Der Zugang zu Bildung ist somit in diesem Verständnis einer der zentralen Faktoren in der Gestal-tung der gesellschaftlichen Rahmenbedingungen, die ein gelingendes Leben befördern können (Nussbaum 2010).

Vor diesem Hintergrund wird im Rahmen des Beitrags empirisch nachgezeichnet, wie sich die Segregation nach sozialer Herkunft in einer mittelgroßen deutschen Stadt (ca. 334.000 Einwohner) un-ter den Bedingungen der freien Schulwahl darstellt. Es wird mit Hilfe von quantitativen Datenanaly-sen den Fragen nachgegangen, welche Familien die Möglichkeit der freien Schulwahl annehmen und welche Veränderungen der Kompositionen in den Schulklassen durch die so entstehenden Wanderungen zwischen den Schulbezirken entstehen. Die Datenbasis für die Untersuchung stellen Angaben zu Wohnbezirken, aufnehmenden Schulen sowie Merkmalen des sozioökonomischen Hintergrundes und der Migrationsgeschichte der Schüler*innen dar und umfassen den kompletten Einschulungsjahrgang des Jahres 2021 (n=3.243).

Es zeigt sich innerhalb der einzelnen Schuleinzugsbereiche des Stadtgebiets eine starke Segregation nach sozialer Herkunft. Auch wenn für alle Kinder die Möglichkeit besteht, Schulen in anderen Stadtbezirken zu besuchen, zeigt sich, dass hauptsächlich Familien mit einem im Vergleich zum Durchschnitt des Wohnbezirks höheren sozialen Status die Möglichkeit nutzen, ihre Kinder auf Schulen eines anderen Schulbezirks anzumelden. Eine besondere Rolle spielen hierbei Schulen, die in nicht-städtischer Trägerschaft sind und ein besonderes pädagogisches Profi aufweisen. Es kön-nen mit Hilfe der Daten Hinweise darauf herausgearbeitet werden, dass schon der Übergang vom Elementar- in dem Primarbereich Grundlagen für irreguläre Bildungsverläufe legt.

Irreguläre Schullaufbahnen durch die Sekundarstufe – Eine Mixed-Methods-Analyse zur Rolle von Klassenwiederholungen im luxemburgischen Bildungssystem

Susanne Backes
Universität Luxemburg

„Pays de redoublement“ – Luxemburg, das Land der Klassenwiederholung, so lautete 2008 die Diagnose einer Forschergruppe, die sich mit verschiedenen Differenzierungsmaßnahmen in heterogenen Lerngruppen befasst hat (Reichert & Dierendonck, 2008). Mehr als 10 Jahre später lässt sich in Luxemburg weiterhin ein hoher Anteil an Klassenwiederholer/innen sowie ein hoher Anteil an Schüler/innen, die älter sind als ihr theoretisches Alter in der jeweiligen Klassenstufe ausmachen (Klapproth et al., 2021; cf. also Valbuena et al., 2020). Dies scheint angesichts vielfach replizierter Forschungsbefunde zur Nicht-Effektivität von Klassenwiederholung in Bezug auf die weitere Kompetenzentwicklung einerseits verwunderlich (Ertel et al., 2021; Klapproth & Schaltz, 2015; Hornung et al., 2021). Andererseits sind stratifizierte Bildungssysteme, zu denen Luxemburg in einem ausgeprägten Maße gehört, anfällig für selektierende Strategien zur Herstellung leistungshomogener Schülergruppen und weisen damit oftmals höhere Wiederholerraten auf.

Klassenwiederholungen sind allerdings nicht nur relevant, wenn es um die Kompetenzentwicklung von Schüler/innen geht, sondern können auch zentrale Bildungsereignisse oder Brüche innerhalb von Bildungsverläufen darstellen, deren Ursachen und Konsequenzen im Kontext weiterer laufbahnbezogener Aspekte zu sehen sind. In diesem Beitrag werden Klassenwiederholungen daher

ähnlich wie der Übergang in die Sekundarstufe als biographische Herausforderung (Koch, 2006) innerhalb von Schulverläufen verstanden. Die Forschungsfragen lauten: Welche Formen irregulärer Bildungsverläufe lassen sich im stratifizierten Bildungssystem Luxemburgs identifizieren? Inwiefern treten Klassenwiederholungen in Verbindung mit weiteren Bildungsereignissen auf? Welche Schülergruppen sind besonders häufig von irregulären Verläufen und Klassenwiederholungen betroffen? Und, wie werden Klassenwiederholungen von Schüler/innen erlebt?

Im Rahmen des Vortrags wird zunächst das luxemburgische Schulsystem hinsichtlich seiner institutionellen Rahmenbedingungen (u. a. dreisprachiges Schulsystem, hohe Stratifizierung) und Schülerpopulation (hohe Heterogenität mit einem Migrantenanteil) charakterisiert. Dem folgt ein theoretischer Abriss der zugrundeliegenden, ungleichheitstheoretischen Ansätze von Boudon (1974) und Bourdieu (1983) sowie Konzepte aus der Jugend- und Transitionsforschung (u. a. Zinnecker 2003), ergänzt um empirische Befunde zu Klassenwiederholungen. Der Diskussion der Resultate wird eine Beschreibung der verwendeten Datensätze vorangestellt. Als Mixed-Methods-Studie angelegt basieren die Analysen auf einem administrativen Panel-Schülerdatensatz (Geburtsjahrgang 1990, Vollerhebung mit n = 3.375) und einer narrativen Interviewstudie mit (ehemaligen) Schüler/innen, die unterschiedlichen Verlaufstypen angehören (Backes, 2018; Backes & Hadjar, 2017).

Die Sekundärdatenanalyse zeigt, dass die Wiederholungsrate im akademischen Track durchschnittlich geringer ist als in den technischen und berufspraktischen Schulformen. Hohe Wiederholungsraten sind insbesondere in den Abschlussklassen festzustellen sowie auch vor der anstehenden Orientierungsentscheidung in der Sekundarstufe des technischen Tracks. Letztere lassen sich teilweise auf strategische Klassenwiederholungen zurückführen, wobei Regressionsmodelle zeigen, dass diese Entscheidung eine sozial-selektive ist. Eine Analyse der Begleitfaktoren, die gemeinsam mit Klassenwiederholungen in der luxemburgischen Sekundarstufe auftreten, zeigt, dass Klassenwiederholungen häufig in Kombination mit einem Schulformwechsel oder Sektionswechsel vorkommen. So ist dann auch zu erklären, dass sehr hohe Anteile an Klassenwiederholungen in abwärtsmobilen Laufbahntypen (80 % mit Wiederholungsereignis) als auch aufwärtsmobilen Laufbahntypen mit Integrationsbewegung (82 %; durchschnittlich 1,6 Wiederholungen je Schüler/in) zu finden sind.

Auch aus den narrativen Interviews geht der kontextspezifische Befund hervor, dass Klassenwiederholungen in Luxemburg als gängiges und nicht notwendigerweise stigmatisierendes Ereignis wahrgenommen werden („Aus meiner Klasse war fast ein Viertel durchgefallen“, IBio10). Die Interviewanalyse verdeutlicht, wie stark die Klassenwiederholung als (kritisches) Lebensereignis in Zusammenhang mit der neuen Peergruppe wahrgenommen und demnach positiv oder negativ erinnert wird (z. B. „Ich bin in eine Klasse gekommen, die mich sofort aufgenommen hat und integriert und alles. Eigentlich war das Sitzenbleiben das Beste, was mir passieren konnte“, IBio13). Auch die eigene Leistungsentwicklung wird von den Jugendlichen u. a. im Lichte der neuen Klassenkomposition interpretiert („Und dann hat man mich in eine Klasse geschickt, wo nur Leute waren, die auch wiederholt haben“, IBio13), sodass neben den betreffenden Lehr-/Lernarrangements auch Fischteicheffekte zum Tragen kommen (Marsh et al., 2007).

Die Erkenntnisse werden auf Basis jüngster Befunde des jährlichen Bildungsmonitorings diskutiert.

Schullaufbahnen von Klassenwiederholenden

Paul Fabian, Katharina Molitor
Technische Universität Dortmund, Germany

Theoretischer Hintergrund

Klassenwiederholungen sind eine national wie international höchst umstrittene Homogenisierungsmaßnahme. In Deutschland haben einige Bundesländer (z.B. Hamburg oder Berlin) das Sitzenbleiben bereits abgeschafft. In Bayern oder Bremen dagegen liegt die Wiederholendenquote über dem Bundesdurchschnitt (vgl. Statistisches Bundesamt 2018). Im internationalen Vergleich ist die Quote der Schülerinnen und Schüler, die im Verlaufe ihrer Bildungskarriere mindestens einmal eine Jahrgangsstufe wiederholt haben, in Deutschland mit ca. 18 % überdurchschnittlich hoch (vgl. Sälzer, et al. 2016). Empirisch ist die Effektivität von Klassenwiederholungen bislang nicht ausreichend belegt. Internationale Studien (Goos et al. 2021) konnten zeigen, dass eine kurzfristige Leistungssteigerung im wiederholten Schuljahr eintrat, welche aber mittel- und langfristig abnahm. Andere Studien zeigten mittel- und langfristig sogar negative Effekte (Hattie 2011). In Deutschland finden sich nur wenige belastbare Studien, die konkrete Aussagen zur Effektivität von Klassenwiederholungen zulassen (z.B. Ehmke et al. 2017, Marsh et al. 2017, Fabian 2020). Ehmke et al. (2017) sowie Fabian (2020) konnten keine positiven Effekte auf die Leistungsentwicklung feststellen. Fabian (2020) konnte zudem zeigen, dass sich keine signifikante Verbesserung von Schulnoten einstellt. Marsh et al. (2017) konnten demgegenüber eine Leistungssteigerung in Mathematik feststellen. Für die Wirkung von Klassenwiederholungen auf den erreichten Schulabschluss konnte Bellenberg (1999, 2012) darstellen, dass eine Klassenwiederholung sehr häufig mit einer Abschulung einhergeht und so die Wahrscheinlichkeit für höhere Abschlüsse geringer wird. Fertig (2004) zeigte mit seinem Forschungsansatz, dass die Wahrscheinlichkeit eines höheren Abschlusses um 48% steigt. Demski und Liegmann (2014) finden in ihrer Analyse nur geringe Unterschiede zwischen Wiederholenden und Nicht-Wiederholenden. Im Kontext der theoretischen Annahmen des Kredentialismus (vgl. Collins 1979, Bills 2003) sind die Befunde von DiPrete et al. (2017) spannend: im internationalen Vergleich kommt in Deutschland Zertifikaten eine besonders große Bedeutung für den Erfolg im Ausbildungs- und Arbeitsmarkt zu. Gleichzeitig sind objektive Kompetenzen und Zertifikate nicht deckungsgleich (vgl. Brändle & Pohlmann 2021). Zusätzlich gibt es Befunde, dass insbesondere niedrig qualifizierte Schülerinnen und Schüler besonders dann erfolgreich in den Ausbildungsmarkt übertreten, wenn sie gute Noten und höhere Aspirationen aufweisen – unabhängig ihrer kognitiven und sozialen Kompetenzen (vgl. Holtmann et al. 2017).

Fragestellung

Vor dem Hintergrund der dargestellten theoretischen Überlegungen und empirischen Erkenntnisse stellt sich die Frage, ob eine Klassenwiederholung die Wahrscheinlichkeit erhöht, ein schulformadäquates Zertifikat zu erreichen oder sogar höhere Zertifikate zu erwerben.

Methode

Datengrundlage bildet die National Educational Panel Study (NEPS, Startkohorte 3; Blossfeld et al.2011). Die Stichprobe besteht aus 6.491 Schülerinnen und Schülern. Die Stichprobengröße reduziert sich bei der Analyse durch den Ausschluss von Kindern, zu denen keine Informationen zum Schulabschluss (CASMIN, vgl. König et al. 1988) vorliegen (N = 4.736). Die fehlenden Werte von Kovariaten wurden mittels multipler Imputation (vgl. R 4.1 (R Core Team 2020), package mice (van Buuren & Groothuis-Oudshoorn 2011), m=50) behandelt. Innerhalb dieser Stichprobe wurden unter Nutzung des Rubin Causal Models (vgl. Rubin 1974) sowie des Propensity Score Matchings (vgl. Rosenbaum & Rubin 1985; Stuart 2010) auf Basis von Informationen von Schüler*innen, Eltern sowie Klassenlehrkräften Propensity Scores berechnet. Daraufhin wurden Klassenwiederholende mit Nicht-Wiederholenden gematcht (1:1 Nearest Neighbour-Matching, Caliper = .3). So können Klassenwiederholende mit jenen Schüler*innen verglichen werden, die ihnen in zentralen Hintergrundmerkmalen ähneln, jedoch regulär versetzt wurden.

Zunächst werden Deskriptiva berichtet und grafisch dargestellt. Danach werden mittels logistischer Regressionen Odds-Ratios für (a) einen schulformadäquaten (Hauptschulabschluss an Hauptschule/Realschulabschluss an Realschule/Abitur an Gymnasien und Gesamtschulen) und (b) für einen höheren Abschluss berichtet.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse deuten darauf hin, dass vergleichbare Klassenwiederholende und Nicht-Wiederholende keine signifikanten Unterschiede in ihren erreichten Schulabschlüssen zeigen Im Einklang mit bestehenden Befunden (vgl. Demski & Liegmann 2014) deuten die Ergebnisse darauf hin, dass die Klassenwiederholung keinen Effekt auf den erreichten Bildungsabschluss hat. Demnach wäre die Maßnahme „Klassenwiederholung“ kritisch zu überdenken.

A-17: Metakognitive Überwachung beim Lernen in Schule und Hochschule: Maßnahmen zur Förderung der Urteilsgenauigkeit

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* T03 R04 D10

Metakognitive Überwachung beim Lernen in Schule und Hochschule: Maßnahmen zur Förderung der Urteilsgenauigkeit

Chair(s): **Julia Waldeyer** (Ruhr-Universität Bochum, Germany)
*Diskutant*in(nen):* **Marion Händel** (Universität Augsburg, Germany)

Akkurate metakognitive Überwachung ist eine fachübergreifende essenzielle Komponente effektiven selbstregulierten Lernens in Schule und Hochschule. Sowohl Theorien als auch empirische Befunde zeigen deutlich, dass Lernende auf Basis ihrer metakognitiven Überwachungsprozesse Folgeentscheidungen zur gezielten Steuerung des Lernverhaltens treffen. Im Speziellen nutzen Lernende ihre metakognitiven Urteile dafür, um über die Dauer, die Art und den Inhalt ihrer Lernaktivitäten zu entscheiden. Allerdings zeigt die Empirie auch, dass Lernende oftmals dazu neigen, ihren Wissensstand nicht akkurat einzuschätzen. So neigen Lernende häufig zur Überschätzung. Diese Fehleinschätzungen können dazu führen, dass Lernende abträgliche Regulationsentscheidungen treffen – beispielsweise, dass sie entscheiden, den Lernprozess zu früh zu beenden. Angesichts des skizzierten Problems wurden in den letzten Jahren in verschiedenen Fachdomänen vielversprechende Maßnahmen zur Förderung akkurater Urteilsgenauigkeit entwickelt und erprobt. Bezüglich dieser Maßnahmen und ihrer spezifischen Wirkmechanismen bestehen allerdings noch offene Fragen. Das Ziel des Symposiums ist es, mithilfe von vier Beiträgen zur Erprobung von Fördermaßnahmen zur Verbesserung der Urteilsgenauigkeit mit Lernmaterialien aus den Disziplin Germanistik, Biologie, Erziehungswissenschaft und Psychologie, zur Klärung dieser Fragen beizutragen. Die Beiträge werden von einer Expertin des Forschungsbereichs zur Förderung metakognitiver Urteilsgenauigkeit diskutiert, die sowohl einen psychologischen als auch einen fachdidaktischen Hintergrund (insbesondere Biologie) aufweist.

In Beitrag 1 wird die Rolle von Emotionen für die Genauigkeit von Selbsteinschätzungen Studierender thematisiert. Im Speziellen wird in den Blick genommen, ob Emotionen beim Lernen aus Texten einen Einfluss auf metakognitive Urteile und deren Genauigkeit haben. Darüber hinaus wird die Frage adressiert, inwiefern eine Reattributionsinstruktion den angenommenen Einfluss der Emotionen mindert. Beitrag 2 beschäftigt sich, ebenfalls im Kontext des Lernens aus Texten, mit dem Einfluss sogenannter Rubrics auf die Genauigkeit metakognitiver Urteile bei Schüler*innen und Studierenden. Zudem werden mögliche Auswirkungen von Rubrics auf Regulationsentscheidungen und die anschließende Aufgabenleistung thematisiert. In Beitrag 3 wird im Kontext des multimedialen Lernens untersucht, wie sich eine explizite Bewertung des eigenen Lernprozesses auf die metakognitive Urteilsgenauigkeit auswirkt. Hier wird abermals in den Blick genommen, welche möglichen Auswirkungen die instruktionale Maßnahme auf Regulationsentscheidungen und auf die anschließende Aufgabenleistung hat. Beitrag 4 schließlich untersucht die Wirkung von externalen Standards bei der Bewertung von Beispielen, die Studierende zu Konzepten aus dem Bereich der Sozialpsychologie generieren. Spezifisch wird die Wirkung von Experten- und Negativbeispielen fokussiert und die Frage adressiert, welche Art von Beispielen – und in welcher Kombination – die Urteilsgenauigkeit optimiert.

Gemeinsam versprechen die vier Beiträge, die jeweils auf experimentellen Studien basieren, bedeutende Fortschritte auf dem Weg zu effektiven Maßnahmen zur Förderung der Urteilsgenauigkeit Lernender in Schule und Hochschule.

Beiträge des Symposiums

Der Einfluss von Hoffnung*slosigkeit und einer Reattributionsinstruktion auf Selbsteinschätzungen des Textverständnisses

Anja Prinz-Weiß
Pädagogische Hochschule Karlsruhe

Theoretischer Hintergrund

Lernende stehen häufig vor der Herausforderung, sich eigenständig Wissen aus Texten aneignen zu müssen. Damit dies gelingt, ist es wichtig, dass die Lernenden ihr eigenes Verständnis akkurat einschätzen, sodass sie adaptive Regulationsprozesse (z.B. erneutes Lesen) eingehen können (z.B. Thiede et al., 2003). Lernende treffen jedoch häufig inakkurate Einschätzungen über ihr Verständnis. Ein Grund dafür ist, dass sie Informationen – sogenannte Cues – (z.B. Vertrautheit mit der Lerndomäne) heranziehen, die nicht prädiktiv für ihr Verständnis sind (z.B. Thiede et al., 2009). Obwohl Emotionen in Lernsituationen stets präsent sind, wurde ihre mögliche Rolle als Cues und inwiefern sie die Genauigkeit von Selbsteinschätzungen beeinflussen, bislang kaum beforscht. Erste Untersuchungen deuten darauf hin, dass Lernende Emotionen und besonders ihre Hoffnung bzw. Hoffnungslosigkeit beim Lernen als Cues verwenden, wobei dies die Genauigkeit der Selbsteinschätzungen beeinträchtigen kann. Eine Limitation dieser ersten Analysen ist jedoch, dass die Lernenden keinen hohen Grad an negativen Emotionen empfanden (Autoren, eingereicht). Weiter scheint es wichtig, instruktionale Methoden ausfindig zu machen, um Lernenden dabei zu helfen, weniger auf ihre Emotionen und stattdessen auf validere Cues zu vertrauen.

Fragestellungen

In dieser Studie wurde untersucht, inwiefern die Emotion Hoffnung*slosigkeit Selbsteinschätzungen des Textverständnisses sowie deren Genauigkeit beeinflusst. Weiterhin wurde beforscht, inwiefern eine Reattributionsinstruktion Lernende dabei unterstützt, ihre Selbsteinschätzungen nicht basierend auf ihren Emotionen zu treffen.

Methode

Insgesamt nahmen 60 Studierende an der Studie mit einem zweifaktoriellen Design teil. Die Probanden wurden zufällig der hoffnungsvollen, hoffnungslosen oder neutralen sowie der Reattributions- oder Kontrollinstruktion zugewiesen. Der hoffnungsvollen Bedingung wurde mitgeteilt, dass Wissen über die Textinhalte wichtig ist, wobei diese gut zu verstehen sind. Die hoffnungslose Bedingung wurde ebenfalls über die Wichtigkeit der Inhalte formiert, wohingegen diese als schwer zu verstehen beschrieben wurden. Der neutralen Bedingung wurde erklärt, dass die Inhalte wichtig sind und sie den Text unvoreingenommen lesen sollen. Anschließend wurde zur Manipulationskontrolle die Hoffnung*slosigkeit der Probanden erfasst. Ziel der Manipulation war es, eine möglichst große Streubreite der Emotion zu bewirken. Als Nächstes lasen die Probanden einen expositorischen Text über Wärmepumpen. Daraufhin wurden sie gebeten, ihre empfundene Hoffnung*slosigkeit während

des Lesens anzugeben. Folgend bekam ein Teil der Probanden die Reattributionsinstruktion. Diese verwies sie explizit darauf, dass Emotionen nicht zwangsläufig ein gültiger Indikator für das erzielte Verständnis darstellen und sie deshalb für ihre Einschätzungen, statt auf diese zu vertrauen, überlegen sollten, wie gut ihr Verständnis der zentralen Textaussagen tatsächlich ist. Der andere Teil bekam die Kontrollinstruktion mit Informationen zu Speed Reading. Die Probanden schätzten dann ihr Verständnis ein (Prädiktion) und bearbeiteten Verständnisfragen. Abschließend gaben die Probanden erneut eine Einschätzung über ihr Verständnis ab (Postdiktion).

Ergebnisse und Diskussion

Durch die Manipulation gab es eine große Streubreite hinsichtlich der empfundenen Hoffnung*slosigkeit während des Lesens, welche als metrische unabhängige Variable verwendet wurde. Die Art der Instruktion diente als dichotome unabhängige Variable. Abhängige Variablen waren die Selbsteinschätzungen (Prä- und Postdiktionen) sowie deren Genauigkeit. Die Ergebnisse zeigten, dass die Hoffnung*slosigkeit sowohl für Prädiktionen, $\beta=.62$, $p<.001$, als auch für Postdiktionen, $\beta=.35$, $p=.010$, als Cue verwendet wurde. Eine höhere Hoffnung ging dabei mit höheren Selbsteinschätzungen einher. Die Reattributionsinstruktion hatte darauf keinen mindernden Einfluss. Hinsichtlich der Genauigkeit der Prädiktionen ergaben sich keine signifikanten Effekte. Bezüglich der Genauigkeit der Postdiktionen gab es einen signifikanten Effekt der Instruktionsart, $\beta=.28$, $p=.034$. Die Reattributionsinstruktion führte zu akkurateren Selbsteinschätzungen. Die Studie zeigt, dass Lernende die Hoffnung*slosigkeit, die sie während des Lernens mit einem Text empfinden, nutzen, um Einschätzungen über ihr Verständnis zu treffen. Eine Reattributionsinstruktion, die verdeutlicht, dass Emotionen kein zuverlässiger Indikator für das erzielte Verständnis sind, scheint nicht ausreichend, um dies zu verhindern. Die Reattributionsinstruktion kann Lernende jedoch dabei unterstützen, akkuratere Postdiktionen zu treffen. Möglicherweise hilft diese den Lernenden, neben ihren Emotionen zusätzlich validere Cues (z.B. Beantwortbarkeit der Verständnisfragen) heranzuziehen.

Rubrics fördern die Akkuratheit von Selbsturteilen, jedoch nicht die Regulation

Rebecca Krebs, Julia Waldeyer, Björn Rothstein, Julian Roelle
Ruhr-Universität Bochum

Theoretischer Hintergrund

Ein häufig eingesetztes Instrument zur Verbesserung der Leistung von Lernenden bei der Bearbeitung verschiedener Aufgaben ist die Gabe von Rubrics (Andrade, 2001). Rubrics zielen darauf ab, Lernende über Bewertungskriterien zu informieren und verschiedene Qualitätsstufen transparent zu machen. Bisherige Untersuchungen zeigen deutlich, dass Rubrics die Leistung verbessern können (Panadero & Jonsson, 2013). Häufig wird die leistungsförderliche Wirkung von Rubrics auf eine Steigerung der Selbsturteilsgenauigkeit zurückgeführt, die wiederum den Weg für effektive Regulation ebnen würde. Trotz der großen Aufmerksamkeit, die Rubrics in der Forschung genossen haben, wurde dieser theoretisch angenommene Mechanismus bisher jedoch kaum geprüft (Andrade, 2019).

Dies gründet unter anderem darauf, dass in den meisten bisherigen Studien Rubrics nicht nur während der Selbstbeurteilung, sondern auch während der Ausführung der Aufgabe selbst bereitgestellt wurden, was es schwierig macht, die Auswirkungen von Rubrics auf die

Selbsturteilsgenauigkeit und die anschließende Regulation zu isolieren. Eine Ausnahme bilden Krebs et al. (2022), die in einem Laborexperiment die Wirkung von Rubrics auf die Selbsturteilsgenauigkeit isolierten und feststellten, dass Rubrics die absolute Genauigkeit verbesserten und den Bias verringerten. Allerdings führten die Autor*innen in ihrer Studie keine Regulationsphase ein und konnten daher nicht prüfen, ob die höhere Selbsturteilsgenauigkeit auch zu verbesserter Regulation des Lernens und daraus resultierend zu verbesserter Leistung führt.

Fragestellung

Vor dem geschilderten Hintergrund verfolgte die vorliegende Studie zwei Ziele. Um die Annahme zu untermauern, dass Rubrics die Selbsturteilsgenauigkeit erhöhen können, sollte die Studie von Krebs et al. (2022) konzeptuell repliziert und somit geprüft werden, ob Rubrics bei der Selbstbeurteilung die absolute Genauigkeit erhöhen und den Bias verringern (Hypothese 1).

Das zweite Ziel war, den Regulationsprozess nach der Selbstbeurteilungsphase zu beleuchten. Insbesondere waren wir an den Auswirkungen von Rubrics auf Regulationsentscheidungen interessiert. Spezifisch sollte geprüft werden, ob Rubrics das Ausmaß erhöhen, in dem Lernende sich dafür entscheiden, Inhalte zu überarbeiten, die sie zuvor am wenigsten erfüllt hatten (Hypothese 2), und ob dies folglich zu einer höheren Aufgabenleistung bei den Lernenden mit Rubrics führt (Hypothese 3).

Methode

Teilnehmende waren N = 118 Schüler:innen der 11. und 12. Jahrgangsstufe sowie Studierende in der Studieneingangsphase (1. und 2. Fachsemester), 85 weiblich, 33 männlich, MAlter = 22.30, SD = 3.44. Die Lernenden lasen einen wissenschaftlichen Artikel und wurden dazu aufgefordert, zu diesem ein Abstract zu schreiben. Danach beurteilten sie ihr Abstract entweder mit oder ohne die Unterstützung einer Rubric. Im Anschluss konnten die Lernenden ein Kriterium auswählen, zu dem sie weitere Instruktion erhalten wollten, und hatten die Möglichkeit, ihr Abstract im Anschluss an die weitere Instruktion zu überarbeiten.

Ergebnisse und Diskussion

Mit Blick auf Hypothese 1 zeigten die Ergebnisse wie erwartet erhöhte absolute Genauigkeit, $t(116) = -1.83$, $p = .035$, $d = 0.33$, sowie verringerten Bias, $t(116) = 1.75$, $p = .041$, $d = 0.32$) in der Gruppe mit Rubrics. In Bezug auf die Regulationsentscheidungen zeigte sich ebenfalls ein statistisch signifikanter Unterschied zwischen den Gruppen, $t(116) = 2.01$, $p = .023$, $d = 0.39$). Entgegen der Annahme entschieden sich jedoch die Lernenden in der Gruppe mit Rubrics weniger häufig dafür, das Kriterium zu verbessern, das sie selbst als am wenigstens erfüllt eingeschätzt hatten. Bezüglich Hypothese 3 zeigten die Ergebnisse, dass die Lernenden ihre Leistung nach der Überarbeitungsphase verbessert hatten, $F(1, 116) = 55.09$, $p < .001$, $\eta p2 = .32$). Jedoch unterschied sich die Verbesserung nicht zwischen den beiden Gruppen, $F(1, 116) = 0.07$, $p = .796$, $\eta p2 < .01$).

Die Ergebnisse zeigen insgesamt, dass der Gebrauch einer Rubric zwar zu verbesserter Selbsturteilsgenauigkeit führt, jedoch nicht in zwangsläufig besseren Regulationsentscheidungen resultiert. Dies stellt die bisher theoretische angenommenen Wirkmechanismus von Rubrics infrage.

Wirkt sich die explizite Bewertung des eigenen Lernprozesses auf Lernverhalten, Lernergebnis und metakognitive Urteilsgenauigkeit aus?

Emely Hoch[1], Katharina Fleig[1], Katharina Scheiter[2]
[1]Leibniz-Institut für Wissensmedien, Tübingen, [2]Universität Potsdam

Theoretischer Hintergrund

Metakognitive Prozesse, wie das Überwachen (monitoring) oder Anpassen (control) des eigenen Lernfortschritts, spielen eine wichtige Rolle, wenn es um die effektive Verarbeitung und das Verstehen von Informationen im Allgemeinen (Nelson & Narens, 1990) oder das Lernen mit Texten und Bildern, d.h. multimediales Lernen, im Besonderen geht (Moreno & Mayer, 2007). Beim Monitoring verfolgen Lernenden ihre kognitiven Prozesse, das bedeutet, sie beobachten, bewerten und reflektieren sie (Nelson & Narens, 1990). Das Ergebnis dieses Überwachungsprozesses wird genutzt, um bewusste Entscheidungen zur gezielten Steuerung des Lernverhaltens zu treffen. Lernende haben jedoch oft Schwierigkeiten, ihren Wissenstand beim multimedialen Lernen genau einzuschätzen, da sie sich beispielsweise auf eine Multimediaheuristik verlassen und ihren Lernstand überschätzen (Serra & Dunlosky, 2010). Generative Prozesse während des Lernens, wie z.B. Erklären oder Abrufübungen, verlangen von Lernenden, dass sie ihr aktuelles Wissen externalisieren, und können daher metakognitive Prozesse auslösen (Fukaya, 2013). Entsprechend zeigen Forschungsbefunde, dass Abrufübungen die metakognitive Urteilsgenauigkeit erhöhen (z.B. Fukaya, 2013; King et al., 1980; Little & McDaniel, 2015) und die Lernleistung verbessern (z.B. Ariel & Karpicke, 2018; Cogliano et al., 2021; Little & McDaniel, 2015). Solche Abrufübungen dienen jedoch nur implizit der Überwachung des Lernprozesses, da diese zwar den aktuellen Wissensstand der Lernenden sichtbar machen, Lernende ihre Lernprozesse aber nicht explizit bewerten.

Fragestellung

In der vorgestellten Studie wurde untersucht, wie sich eine zusätzliche, explizite Bewertung des eigenen Lernprozesses auf multimediales Lernen und die metakognitive Urteilsgenauigkeit auswirkt.

Methode

Die Teilnehmenden lernten 10 Seiten multimediales Material zum Thema Mitose. Nach sieben dieser Seiten erfolgte eine Abrufübung, d.h. die Teilnehmenden sollten aufschreiben, was sie von dem Lerninhalt der zuvor präsentierten Seite erinnerten. Nach dem Lernen gaben sie eine Einschätzung, inwiefern sie Fragen zu den gelernten Inhalten richtig beantworten könnten (metakognitives Urteil). Darüber hinaus wurde die Lernleistung gemessen (offene Wiedergabefrage, 29 Wahr/Falsch-Fragen). Alle Teilnehmenden wurden also gebeten, die Inhalte während des Lernens abzurufen, und verfügten somit über Informationen über ihren Lernstand; allerdings wurde nur eine Gruppe gebeten, ihren Lernprozess explizit zu bewerten: die Hälfte der Teilnehmenden wurde jeweils vor der Abrufübung zusätzlich gebeten, ihren Lernprozess im Hinblick darauf zu bewerten, wie sicher sie seien die Inhalte vollständig und korrekt gelernt zu haben, wie viel Anstrengung sie investiert und benötigt haben, sowie wie schwierig es war, die Inhalte zu lernen. Es wurde erwartet, dass die explizite Bewertung des eigenen Lernprozesses zusätzlich zur Abrufübung das Lernverhalten (Lernzeit, Antwortlänge bei der Abrufübung) beeinflussen sowie die Lernleistung und die metakognitive Urteilsgenauigkeit (Diskrepanz metakognitives Urteil und tatsächliche Lernleistung) verbessern würde.

Ergebnisse und Diskussion

Die Daten von 143 Teilnehmenden (100 weiblich, Alter: M = 23.8, SD = 4.2) wurden analysiert, die an der online durchgeführten, etwa 45 min dauernden Studie teilnahmen. Entgegen der Annahmen konnten keine Unterschiede in der Lernleistung oder der Lernzeit für die Gruppen ohne und mit expliziter Bewertung des Lernprozesses gezeigt werden, alle $p > .05$. Jedoch unterschieden sich die Gruppen in der metakognitiven Urteilsgenauigkeit: Teilnehmende, die ihren Lernprozess explizit bewerteten, zeigten eine genaue Einschätzung ihres Lernstandes, unter- oder überschätzten ihren Lernstand also nicht signifikant, $p > .1$, während Teilnehmende ohne explizite Bewertung ihren Lernstand signifikant überschätzten, $p < .01$. Zudem schrieben Teilnehmende, die ihren Lernprozess explizit bewerteten, signifikant längere Antworten bei den Abrufübungen, $p < .05$, was als Indikator für die Regulation des Lernprozesses interpretiert werden kann. Zudem zeigte sich ein signifikant indirekter Effekt von expliziter Bewertung über Antwortlänge bei den Abrufübungen auf die gezeigte Lernleistung, $p < .01$.

Die Ergebnisse lassen vermuten, dass die explizite Bewertung des Lernprozesses den Lernenden hilft ihren Lernstand akkurater einzuschätzen, jedoch nicht unbedingt zu besseren Lernleistungen führt. Die explizite Bewertung des Lernprozesses wirkt sich jedoch über eine Anpassung des Lernprozesses indirekt auf die Lernleistung aus.

Erhöhen externale Standards die Urteilsgenauigkeit von Lernenden bei der Beurteilung selbst-generierter Beispiele?

Linda Froese, Julian Roelle
Ruhr-Universität Bochum

Theoretischer Hintergrund

Das Generieren eigener Beispiele ist eine effektive Lernaktivität zur Aneignung von konzeptuellem Wissen (Rawson & Dunlosky, 2016). Um positive Lernergebnisse zu erzielen, sollten diese Beispiele von hoher Qualität sein. Daher sollten Lernende in der Lage sein, die Qualität ihrer Beispiele akkurat einzuschätzen, um entscheiden zu können, wann sie weitere Anstrengungen in diese Lernaktivität investieren sollten und wann nicht. Leider schätzen Lernende die Qualität ihrer Beispiele allerdings oftmals nicht sehr akkurat ein (Froese & Roelle, 2022; Zamary et al., 2016).

Ein wirksames Mittel zur Förderung der Urteilsgenauigkeit bei der Bewertung selbst-generierter Produkte ist die Bereitstellung externaler Standards. Dieser Ansatz wurde ursprünglich für den Einsatz in Abrufübungsaufgaben entwickelt (Lipko et al., 2009), hat sich aber auch innerhalb anderer Lernkontexte bewährt. In einer Studie von Froese und Roelle (2022) führte die Darbietung von Standards (hier: Expertenbeispiele) ebenfalls zu einer Verbesserung der Urteilsgenauigkeit. Allerdings überschätzten die Lernenden dennoch ihre Beispiele von geringer Qualität bedeutend. Ein mögliches Mittel dem entgegenzuwirken könnte darin bestehen, den Lernenden zusätzlich Negativbeispiele, d.h. Beispiele von geringer Qualität, für den Vergleich mit ihren Beispielen zur Verfügung zu stellen. Dies könnte den Lernenden helfen, Parallelen zu den eigenen Beispielen mit geringerer Qualität zu erkennen und angemessenere Urteile zu treffen.

Fragestellung

Vor diesem Hintergrund haben wir eine Studie konzipiert, um erstens die Wirkung von Expertenbeispielen zu replizieren und zweitens zu testen, ob Negativbeispiele, die falsche Illustrationen der jeweiligen Lerninhalte enthalten, einen Mehrwert für die Urteilsgenauigkeit mit sich bringen. Konkret interessierte uns, ob nicht nur Expertenbeispiele, sondern auch Negativbeispiele die Urteilsgenauigkeit verbessern und ob diese Effekte additiv oder sogar überadditiv sind und daher beide Arten von Beispielen zusammen am besten funktionieren.

Methode

Um die Auswirkungen von Experten- und Negativbeispielen zu testen, haben wir eine Studie mit 128 Universitätsstudierenden durchgeführt. Zunächst absolvierten alle Lernenden einen Vortest und lernten acht Konzepte aus dem Bereich der Sozialpsychologie mithilfe eines Textes und Definitionen (angelehnt an das Material von Zamary et al., 2016). Die Teilnehmenden wurden dann aufgefordert, ein Beispiel nach dem anderen für jedes Konzept zu generieren und diese mit 0, 0.5 oder einem Punkt zu bewerten. Während dieses Beurteilungsverfahrens variierten wir in einem 2×2-faktoriellen Design, ob die Lernenden Expertenbeispielstandards (mit vs. ohne) und Negativbeispielstandards (mit vs. ohne) zum Vergleich erhielten. Mithilfe von Lernenden- und Expertenbewertungen wurden als abhängige Variablen zwei Maße der Urteilsgenauigkeit (Bias und absolute Genauigkeit) berechnet.

Ergebnisse und Diskussion

Hinsichtlich des Bias (XLernendenurteil - XExpertenurteil) fanden wir einen statistisch signifikanten Haupteffekt des Faktors Expertenbeispiele, $F(1, 123) = 3.96$, $p = .049$, $\eta p^2 = .03$, der auf einen geringeren Bias in den Gruppen mit Expertenbeispielen hindeutete, sowie einen statistisch signifikanten Effekt des Faktors Negativbeispiele, $F(1, 123) = 4.10$, $p = .045$, $\eta p2 = .03$, der einen geringeren Bias in den Gruppen mit Negativbeispielen signalisierte. Darüber hinaus zeigte sich ein statistisch signifikanter Interaktionseffekt, $F(1, 123) = 5.54$, $p = .020$, $\eta p2 = .04$. Die Interaktion zeigte, dass Expertenbeispiele für die Lernenden, die keine Negativbeispiele erhielten, wirksam waren, aber keinen zusätzlichen Nutzen für die Lernenden hatten, die Negativbeispiele erhielten.

Da sich Unter- und Überschätzungen in den Bias-Werten gegenseitig aufheben können, berechneten wir zusätzlich die absolute Genauigkeit (|XLernendenurteil - XExpertenurteil|). Die ANOVA zeigte einen statistisch signifikanten Effekt, $F(1, 123) = 8.50$, $p = .004$, $\eta p2 = .06$, nur für den Faktor Expertenbeispiele. Die Lernenden der Gruppe mit Expertenbeispielen zeigten eine höhere absolute Genauigkeit als die Lernenden ohne Expertenbeispiele.

Die Ergebnisse deuten darauf hin, dass insbesondere die Gabe von Expertenbeispielen die Urteilsgenauigkeit von Studierenden bei der Bewertung selbst-generierter Beispiele erhöhen kann, während die Effekte der Negativbeispiele uneindeutig sind. Die Kombination beider Standards zeigte keinen Mehrwert. Weitere Studien sind notwendig, um die Wirkmechanismen hinter den Standards zu beleuchten.

A-18: Professionelle Kompetenz von Fachkräften im Elementarbereich – Gemeinsamkeiten und Spezifika in den Bildungsbereichen Mathematik, Naturwissenschaften und Technik

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* R11 T00 D01

Professionelle Kompetenz von Fachkräften im Elementarbereich – Gemeinsamkeiten und Spezifika in den Bildungsbereichen Mathematik, Naturwissenschaften und Technik

Chair(s): **Simone Dunekacke** (Freie Universität Berlin, Deutschland), **Elisa Oppermann** (Universität Bamberg)
*Diskutant*in(nen):* **Ilonca Hardy** (Goethe Universität Frankfurt am Main)

Kompetenzen in Mathematik, Informatik, Naturwissenschaften und Technik, sog. MINT-Bereiche, gelten als zentral für die gesellschaftliche Teilhabe (Falloon et al., 2020; OECD, 2018). MINT-Bereiche werden deswegen auch in der der Frühpädagogik als relevante Bildungsbereiche für Kinder benannt (KMK, 2004).

Als ausschlaggebend für die Umsetzung der MINT-Bildungsbereiche in der frühen Bildung gelten die professionellen Kompetenzen pädagogischer Fachkräfte. So kommt der Lernunterstützung der Kinder durch die Fachkräfte in Spielsituationen oder Angeboten eine besondere Bedeutung zu (u.a. Weber et al., 2020; Hauser et al., 2014), woraus sich Fragen nach der Bedeutung bereichsspezifischer professioneller Kompetenzen der Fachkräfte ergeben. Dabei können Unterschiede zwischen den Bereichen identifiziert werden. Diese betreffen die Inhalte bzw. Denk- und Arbeitsweisen des Faches, beispielsweise ein empirisch-erklärendes Vorgehen gegenüber einem vom konkreten Problem ausgehendem Handeln (Bergner et al., 2018). Gemeinsam ist den vier Bereichen, dass sie die Förderung von Interesse, Motivation und Selbstwirksamkeitserwartungen der Kinder anregen sollen, was durch die Auseinandersetzung mit grundlegenden Denk- und Arbeitsweisen bzw. inhaltlichen Ideen erfolgt (Bergner et al., 2018; Benz et al., 2017; Graube et al., 2015; Anders et al., 2013).

Forschung zur professionellen Kompetenz von frühpädagogischen Fachkräften hat in den vergangenen Jahren national und international stark an Bedeutung zugenommen (u.a. Dunekacke & Barenthien, 2021; Bruns et al., 2021). Allgemeine Modelle professioneller Kompetenz (Blömeke et al., 2015; Fröhlich-Gildhoff et al., 2011) haben dabei die bereichsspezifisch untersuchten Konstrukte, beispielsweise Wissen, Überzeugungen, situationsspezifische Fertigkeiten, stark geprägt. Mögliche bereichsspezifische aber auch generische Aspekte von MINT-Kompetenzen wurden dabei bislang nicht untersucht oder diskutiert. Solche Synergien könnten für die zukünftige Forschung sowie die Aus- und Weiterbildung von Fachkräften hilfreich sein. Im Rahmen des Symposiums sollen daher Perspektiven auf die professionelle Kompetenz von frühpädagogischen Fachkräften aus der frühen mathematischen, naturwissenschaftlichen und technischen Bildung mit Blick auf Gemeinsamkeiten und Spezifika diskutiert werden.

Im ersten Beitrag wird die Wirksamkeit einer Intervention zur Förderung mathematikbezogener situationsspezifischer Fertigkeiten von angehenden Erzieher*innen im Bereich Mathematik untersucht. Der zweite Beitrag untersucht die Prozessqualität von naturwissenschaftlichen Bildungsangeboten, die berufstätige Fachkräfte anbieten und inwieweit diese mit dem fachdidaktischen Wissen der Fachkräfte zusammenhängt. Der dritte Beitrag untersucht die Nutzung digitaler Medien zur Teamzusammenarbeit. Es zeigt sich, dass neben der technischen Ausstattung,

den Überzeugungen der Fachkräfte und ihrer technischen Selbstwirksamkeitserwartung insbesondere die technische Unterstützung bedeutsam ist.

Die drei Beiträge nehmen damit verschiedene MINT-Bildungsbereiche, verschiedene Aspekte professioneller Kompetenz sowie verschiedene Ansatzpunkte für die Weiterentwicklung der Kompetenzen (u.a. Individuum vs. Team) in den Blick. Die Beiträge werden anschließend von Ilonca Hardy integrativ diskutiert.

Beiträge des Symposiums

Mathematische professionelle Kompetenz von angehenden Erzieher*innen – Der Einfluss von spezifischen Lerngelegenheiten auf die Situationswahrnehmung und Handlungsplanung

Lisa Starcke[1], Julia Barenthien[2], Aiso Heinze[3], Mirjam Steffensky[2]
[1]Freie Universität Berlin, [2]Universität Hamburg, [3]IPN Kiel

Theoretischer Hintergrund

Für die Weiterentwicklung erster mathematischer Konzepte, sind Kinder auf anregende Interaktionen mit frühpädagogischen Fachkräften angewiesen (Lehrl et al., 2016). Für die Umsetzung ist die bereichsspezifische professionelle Kompetenz der frühpädagogischen Fachkräfte ausschlaggebend (Gasteiger & Benz, 2018). Sie ist ein mehrdimensionales Konstrukt, das sich neben den Dispositionen, affektiv-motivationalen Aspekten und der Performanz auch aus den situationsspezifischen Fertigkeiten zusammensetzt (Blömeke et al., 2015; Fröhlich-Gildhoff et al., 2011). Zu diesen handlungsnahen Facetten der professionellen Kompetenz gehören die Situationswahrnehmung (kurz: PERC von engl. perception) und Handlungsplanung (kurz: ACT von engl. action planning). PERC beinhaltet die Wahrnehmung von potentiellen mathematischen Lernsituationen im Kita-Alltag, sowie das Erkennen von Situationen in denen der kindliche Fokus auf der Mathematik liegt. ACT hingegen umfasst das Planen von Handlungen und damit einhergehend die Entscheidung für eine anregende Interaktion (Blömeke et al., 2015; Gasteiger & Benz, 2018).

Es wird angenommen, dass die situationsspezifischen Fertigkeiten den Zusammenhang von Dispositionen und Performanz mediieren (Dunekacke et al., 2015; McCray & Chen, 2012; Oppermann et al., 2016). Weiterhin zeigen Befunde, dass beide Konstrukte bei (angehenden) Erzieher*innen stark miteinander zusammenhängen (Dunekacke et al., 2015; Dunekacke et al., 2016; Lindmeier et al., 2020). Die gleichen Studien geben Hinweise, dass PERC auch als Voraussetzung für ACT interpretiert werden könnte.

Die Rolle der situationsspezifischen Fertigkeiten für die Performanz und der geringe Raum den die mathematische frühe Bildung in Deutschland in den Curricula der Erzieher*innenausbildung einnimmt (Blömeke et al., 2015; Janssen, 2010), verdeutlichen die Relevanz der Identifizierung kurzer, wirksamer mathematischer Lerngelegenheiten für angehende Erzieher*innen. Ob ACT und PERC durch spezifische Lerngelegenheiten gefördert werden können, ist bisher nicht erforscht.

Forschungsfrage

Wie wirkt die bereichsspezifische Ausbildung von Erzieher*innen am besten auf die situationsspezifischen Fertigkeiten?

Methode

Im Rahmen dieser Studie wurden solche Lerngelegenheiten für die frühe mathematische und naturwissenschaftliche Bildung konzipiert und mit Hilfe eines experimentellen, längsschnittlichen Studiendesigns evaluiert. Dabei werden drei Treatmentgruppen und eine Kontrollgruppe unterschieden. Alle Treatmentgruppen erhielten Lerngelegenheiten zum mathematikdidaktischen Wissen und zu situationsspezifischen Fertigkeiten im Umfang von acht Stunden, aufgeteilt auf zwei Tage. Abhängig von der Gruppe, erhielten sie am dritten Tag vier weitere Stunden mit unterschiedlichen Inhalten (Gruppe „best practice“ = situationsspezifische Fertigkeiten im Bereich Naturwissenschaften; Gruppe „transfer“ = Hinweise zum Transfer zwischen den Bereichen Mathematik und Naturwissenschaften, sowie Wissen im Bereich Medien (um die Interventionszeit gleich zu halten); Gruppe „business as usual“ = Wissen im Bereich Medien (um die Interventionszeit gleich zu halten).

Zur Erfassung der mathematischen situationsspezifischen Fertigkeiten, wurde ein Bildvignettentest, der MSSPS-Test (Starcke et al., under review) eingesetzt. An der Studie nahmen zum ersten Messzeitpunkt insgesamt n= 234 Teilnehmende an neun Fachschulen in Berlin und Brandenburg teil.

Die Reliabilitäten für PERC (acht Items) bzw. ACT (sechs Items) lagen im Prä-Test bei einem Cronbach`s Alpha von α = .73 / .53 und beim Post-Test bei α = .70 / .58.

Ergebnisse und Diskussion

Vorläufige Analysen mittels multivariater Varianzanalysen für den Prä- und Post-Test weisen auf keine signifikanten Veränderungen von PERC und ACT hin. Lediglich für die „best-practice“- Gruppe, zeigte sich ein leichter, nicht signifikanter, Zuwachs. Es sind weiterführende Analysen geplant, die im Rahmen des Vortrags vorgestellt und diskutiert werden sollen.

Das Fehlen signifikanter Effekte könnte auf den geringen zeitlichen Umfang der konzipierten Lerngelegenheiten zurückgeführt werden. Die leichte Zunahme an Fertigkeiten in der „best-practice“-Gruppe könnte diese Annahme unterstützen. Sie erhielt am dritten Tag der Interventionen eine vierstündige Lerngelegenheit zu den situationsspezifischen Fertigkeiten im Bereich Naturwissenschaften. Eine intensivere Auseinandersetzung mit diesen Kompetenzfacetten, scheint also förderlich für deren Ausprägung sein.

Innerhalb des Vortrages soll neben den weiterführenden Analysen auch die Inhaltsspezifität der situationsspezifischen Fertigkeiten angehender frühpädagogischer Fachkräfte diskutiert werden.

Prozessqualität von naturwissenschaftlichen Bildungsangeboten in der Kita und die Bedeutung des fachdidaktischen Wissens für die Prozessqualität

Julia Barenthien[1], Yvonne Anders[2], Mirjam Steffensky[1]
[1]Universität Hamburg, [2]Universität Bamberg

Theoretischer Hintergrund

Naturwissenschaften sind in vielen Ländern ein wichtiger Bildungsbereich für die Kita (OECD, 2017). Die Entwicklung naturwissenschaftlicher Kompetenz junger Kinder hängt dabei neben individuellen Voraussetzungen stark von naturwissenschaftlichen Bildungsangeboten in der Kita ab (vgl. Angebot-Nutzungs-Modell nach Roux & Tietze, 2007). Als zentral für die Wirkung der Bildungsangebote wird dabei die Prozessqualität angesehen, die sich auf die Qualität der Interaktionen zwischen Kind und

Fachkraft bezieht. Für die Gestaltung qualitativ hochwertiger domänenspezifischer Bildungsangebote mit z.B. anregenden Interaktionen werden domänenspezifische Aspekte der professionellen Kompetenz der Fachkräfte wie fachdidaktisches Wissen als wichtig angesehen (Siraj-Blatchford et al., 2002). Bisher ist nur wenig über die Prozessqualität von naturwissenschaftlichen Angeboten sowie die Bedeutung des fachdidaktischen Wissens für die Prozessqualität bei naturwissenschaftlichen Bildungsangeboten bekannt. Die wenigen vorliegenden Befunde weisen auf eine geringe Prozessqualität durch die Fachkräfte (Areljung, 2018) sowie positive Zusammenhänge zwischen dem fachdidaktischen Wissen und der Prozessqualität (Gropen et al., 2017) hin. Inwiefern diese Befunde auf Deutschland übertragbar sind, ist auf Grund von Unterschieden in der Ausbildung von Fachkräften sowie der weit verbreiteten Situationsorientierung in den Kitas unklar.

Fragestellung

Ziel dieses Beitrages ist es, zu prüfen, (1) wie die Prozessqualität bei naturwissenschaftlichen Bildungsangeboten und (2) das naturwissenschaftsspezifische fachdidaktische Wissen pädagogischer Fachkräfte ausgeprägt ist und (3) ob und wie das fachdidaktische Wissen mit der Prozessqualität von naturwissenschaftlichen Bildungsangeboten zusammenhängen.

Methode

Im Rahmen des NaQua-Projekts liegen erste vorläufige Daten für eine Teilstichprobe von N=30 Fachkräften in Deutschland vor (93.3% weiblich; MAlter=36.10 Jahre; SDAlter=10.23). Zur Beantwortung der Fragestellungen wurden die Fachkräfte bei der Durchführung eines naturwissenschaftlichen Bildungsangebots videographiert. Diese Videos wurden durch zertifizierte Beobachtende mit dem standardisierten Beobachtungsrating CLASS Pre-K (Pianta et al., 2008) hinsichtlich der Prozessqualität auf einer 7er-Skala geratet. Dabei werden die Domänen emotionale Unterstützung, Gruppenführung und instruktionale Unterstützung durch verschiedene Dimensionen erhoben. Das fachdidaktische Wissen wurde mit einem Testinstrument mit 26 Items erfasst und erzielte eine akzeptable Reliabilität (α=.714). Die Items zum fachdidaktischen Wissen beziehen sich auf Kognitionen (z.B. Kindervorstellungen) sowie auf Instruktionsstrategien (z.B. geeignete Aktivitäten für die Erarbeitung von bestimmten Inhalten).

Ergebnisse

Erste Ergebnisse zeigen im Durchschnitt eine im niedrigen mittleren Qualitätsbereich anzusiedelnde instruktionale Unterstützung (M=3.49; SD=1.00; Min=1.67, Max=5.00) sowie eine hohe emotionale Unterstützung (M=5.97; SD=0.63; Min=4.75, Max=7.00) und Gruppenführung (M=5.53; SD=0.86; Min=3.33, Max=6.67) bei den Bildungsangeboten. Die Fachkräfte erzielten im Mittel 9.7 von 26 Punkten im Test zum fachdidaktischen Wissen (SDFDW=3.61; Min=4.00; Max=16.00). Ergebnisse der Regressionsanalyse, bei der in MPlus eine Standardfehlerkorrektur mit Type Complex zur Berücksichtigung der Mehrebenenstruktur durchgeführt wurde, zeigen, dass das naturwissenschaftsspezifische fachdidaktische Wissen prädiktiv für die Qualität der emotionalen Unterstützung (β=.373; SE=.153;p<.05), instruktionalen Unterstützung (β=.419; SE=.145;p<.01) sowie Gruppenführung (β=.432; SE=.113;p<.01) der naturwissenschaftlichen Bildungsangebote ist.

Diskussion

Die beobachtete Qualität der instruktionalen Unterstützung weist konsistent mit bisherigen Studien (Kauertz & Gierl, 2014) eine große Varianz auf und ist im Mittel im niedrigen mittleren Qualitätsbereich der CLASS-Skala zu verorten (Pianta et al., 2008). Das im niedrigen Qualitätsbereich einzuordnende Minima weist auf den Bedarf der Qualitätsentwicklung hin. Auch die sehr unterschiedlichen Punktzahlen in dem Test zum fachdidaktischen Wissen deuten auf große Unterschiede zwischen den Fachkräften, was sich wie in anderen Studien möglicherweise durch unterschiedliche Lerngelegenheiten zu früher naturwissenschaftlicher Bildung in schulischer Bildung sowie Aus- und Fortbildung erklären lässt (Barenthien et al., 2020). Das fachdidaktische Wissen erweist sich erwartungsgemäß als prädiktiv für die instruktionale Unterstützung. Hingegen sind die positiven Zusammenhänge zur emotionalen Unterstützung sowie Gruppenführung überraschend. Vertiefte Analyse zeigen allerdings nur Zusammenhänge zu den Dimensionen der emotionalen Unterstützung und Gruppenführung, die u.a. auf die Wahrnehmung von akademischen Bedürfnissen der Kinder oder das Aufrechterhalten ihres Interesses bei den Aktivitäten abzielen.

Digitale Teamzusammenarbeit in Kitas – Eine Untersuchung mittels des WST-Models

Sebastian Then, Matthias Borgstede, Yvonne Anders
Universität Bamberg

Digitale Medien sind integraler Bestandteil der Lebenswelt der meisten Kinder in Deutschland (MFS, 2017). Im Zuge der Digitalisierung der Gesellschaft macht diese Entwicklung auch vor Kindertageseinrichtungen nicht halt. Außerdem stellen die im Rahmen der Covid-19 Pandemie beschlossenen Schließungen von Kindertageseinrichtungen (Kitas) pädagogische Fachkräfte vor neue Herausforderungen, die häufig digitale Lösungskonzepte erfordern, beispielsweise in der Zusammenarbeit mit Familien (Cohen et al., 2021) oder dem Kontakt mit Kindern (Gric et al., 2022).

Vor diesem Hintergrund verändert sich auch die Teamzusammenarbeit innerhalb der Kindertageseinrichtungen, insbesondere im Hinblick auf digitale Kommunikationsformen. Für die Kommunikation im Team steht oft kaum Zeit zur Verfügung. Informationen von Fort- und Weiterbildungen einzelner Teammitglieder sollten ans Team weitergegeben werden und die Effizienz von Teamsitzungen und Besprechungen hängt von deren guter Planung, einer transparenten, zuverlässigen Dokumentation und ausreichend Zeit für das Miteinander als Team ab. Die digitale Teamzusammenarbeit ermöglicht, Informationen orts- und zeitunabhängig weiterzugeben und so mehr Teammitglieder in kürzerer Zeit zu erreichen (Leopold & Ullman, 2018). Durch eine verbesserte, digital unterstütze, Teamkommunikation können Informationsunsicherheiten vermieden werden, beispielsweise während des Wechsels zwischen den Schichten oder durch Sammeln von Informationen und Quellen.

Ein mögliches Modell zur Erklärung der Nutzung digitaler Medien, ist das Will-Skill-Tool (WST) Modell von Knezek und Christensen (2016). Die zentralen Bedingungsfaktoren für digitale Mediennutzung sind laut WST die Einstellung, die Fähigkeit (Kompetenz und Selbstwirksamkeitserwartung) und die vorhandene Ausstattung. Bislang gibt es allerdings kaum Forschungsarbeiten zur digitalen Teamzusammenarbeit von pädagogischen Fachkräften in Kitas.

Der vorliegende Beitrag soll dazu beitragen, diese Forschungslücke zu schließen. Im Fokus stehen dabei folgende Fragen: (a) Wie sind frühpädagogische Fachkräfte gegenüber der Nutzung digitaler

Medien zur Teamzusammenarbeit eingestellt und wie nutzen sie diese? (b) Inwiefern kann das WST die Nutzung digitaler Medien in der Teamzusammenarbeit erklären? (c) Wie wichtig für die Nutzung digitaler Medien in der Teamzusammenarbeit ist die Ausstattung, insbesondere die technische Unterstützung der pädagogischen Fachkräfte bei der Nutzung?

Zur Beantwortung der Forschungsfragen wurden Daten aus einer Onlinebefragung (Oktober 2022 bis Dezember 2022) von frühpädagogischen Fachkräften im Rahmen der Evaluation des Bundesprogrammes „Sprach-Kitas“ genutzt. Insgesamt nahmen 886 frühpädagogische Fachkräfte aus 160 Kitas an der Befragung teil. Sie wurden dabei zur Ihrer Nutzungshäufigkeit digitaler Medien, zur Teamzusammenarbeit (Beispielitem: „Mitteilen von Informationen im Team“, $\alpha = .84$), zu ihren Überzeugungen bezüglich digitaler Medien (Beispielitem: „Digitale Medien sind ein guter Ersatz, wenn man sich nicht persönlich treffen kann.“, $\alpha = .77$), zur technischen Ausstattung und zu ihrer technischen Selbstwirksamkeitserwartung (Beispielitem: „Materialien und Infos mit Familien über andere Wege als E-Mail (z.B. Messenger, Dropbox) teilen.“, $\alpha = .76$) befragt. Es wurden Mehrebenenmodelle unter Berücksichtigung der Kita- und Fachkraftmerkmale berechnet.

Die Ergebnisse zeigen, dass die frühpädagogischen Fachkräfte gegenüber der Nutzung digitaler Medien in der Teamzusammenarbeit eher positiv eingestellt sind und sie diese durchschnittlich einmal im Monat nutzen. Die Teilnahme an Fortbildungen geht mit einer verstärkten Nutzung digitaler Medien in der Teamzusammenarbeit einher ($\beta = .13$, $p < .00$). Auch stehen die Einstellungen ($\beta = .12$, $p = .01$), die Ausstattung ($\beta = .09$, $p = .04$), der technische Support ($\beta = .21$, $p < .00$) und die technische Selbstwirksamkeitserwartung ($\beta = .09$, $p = .04$) in positivem Zusammenhang mit der Nutzung digitaler Medien zur Teamzusammenarbeit. Die Modelle erklären 9% der Varianz innerhalb der Teams ($p < .00$) und 8% der Varianz zwischen den Teams ($p = .13$) auf. Keine Zusammenhänge zeigen sich u. a. zwischen dem Geschlecht, dem Alter der frühpädagogischen Fachkräfte sowie der Einrichtungsgröße und der Nutzung digitaler Medien zur Teamzusammenarbeit.

Es zeigt sich, dass zur Nutzung digitaler Medien in der Teamzusammenarbeit sowohl die Überzeugung, die Ausstattung als auch die technische Selbstwirksamkeitserwartung relevant sind. Dabei scheint die technische Unterstützung wichtiger zu sein als das reine Vorhandensein von Computern.

A-19: Medienbezogene Lehrkompetenz: Messung, Wirkung, Förderung

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* R11 T00 D03

Medienbezogene Lehrkompetenz: Messung, Wirkung, Förderung

Chair(s): **Karsten Stegmann** (Universität Passau), **Frank Fischer** (Ludwig-Maximilians-Universität München)
*Diskutant*in(nen):* **Ulrike Cress** (Leibniz-Institut für Wissensmedien)

Die Digitalisierung des Unterrichts in der Schule stellt große Herausforderungen an das Bildungssystem. Während theoretische Modelle und empirische Befunde aufzeigen, wie digitale Medien positive Effekte auf das Lernen in der Schule haben können (vgl. Stegmann, 2022), werden bei Fragen zur Umsetzung eines flächendeckenden digitalen und digital unterstützten Lernens und Unterrichtens in der Schule beträchtliche Forschungsbedarfe deutlich.

Einen zentralen Dreh- und Angelpunkt stellt dabei die medienbezogene Lehrkompetenz dar. Aus dem verbreiteten TPACK-Modell (Koehler & Mishra, 2009) lässt sich ableiten, dass die medienbezogene Lehrkompetenz eine Schnittmenge aus technologischen, pädagogischen und inhaltlichen Kompetenzen darstellt. Das TPACK-Modell definiert die einzelnen Kompetenzbereiche sowie deren Schnittmengen jedoch nicht. Welche Kompetenzen-Kombinationen eine hohe medienbezogene Lehrkompetenz darstellen oder ob es evtl. verschiedene, vergleichbar effektive, Kombinationen geben könnte, bleibt offen.

Diese Defizite ließen sich z. T. mit empirischer Forschung schließen. Aktuell fehlen jedoch noch objektive Instrumente, die auch die Schnittmengen auf technologischem und fachlichem Wissen (TCK, TPCK) abdecken. Offen ist darüber hinaus generell die Validität des Konstrukts „medienbezogene Lehrkompetenz“. Bisherige Modelle wie z. B. das DigiCompEdu (Redecker, 2017) definieren die medienbezogene Lehrkompetenz normativ als die Schnittmenge aus technologischen, pädagogischen und inhaltlichen Kompetenzen. Belastbare empirische Befunde über den tatsächlichen Zusammenhang zwischen der medienbezogenen Lehrkompetenz und dem Kompetenzerwerb der Lernenden oder zumindest der Qualität digital gestützten Unterrichts fehlen noch. Damit fehlen jedoch die Grundlagen für eine an Evidenzen orientierte Lehrkräftebildung in Bezug auf medienbezogene Lehrkompetenz.

Zudem stellt sich die Frage, wie Lehrkräfte in einer sich rasch entwickelnden digitalen Welt dabei unterstützt werden können, medienbezogene Lehrkompetenzen zu erwerben und aktuell zu halten. Professionelle Lerngemeinschaften werden als ein Weg gesehen, medienbezogene Lehrkompetenzen als Teil der Schulentwicklung zu fördern. (Harris, & Sass, 2011; Gräsel & Fussangel, 2010). Offen ist jedoch, wie die Implementation solcher Professioneller Lerngemeinschaften flächendeckend mit den vorhandenen Ressourcen erfolgreich umgesetzt werden kann.

Im Rahmen dieses Symposiums stellen verschiedene Teilprojekte eines Projekts erste Befunde zur Messung, Wirkung und Förderung medienbezogener Lehrkompetenzen von/bei Lehrkräften vor. Im Beitrag von (1) Autoren1 stellen Instrumente zur objektiven Messung von TCK und TPCK vor. Im Beitrag von (2) Autoren2 wird analysiert, inwieweit medienbezogene Lehrkompetenzen die Unterrichtsqualität und die Medienkompetenz (ICT Literacy) von Schülerinnen und Schülern vorhersagen können. Die Beiträge von (3) Autoren3 sowie (4) Autoren4 der Frage, inwieweit durch

die Implementation von Professionellen Lerngemeinschaften an einzelnen Schulen bzw. schulübergreifend die für die Digitalisierung des Unterrichts an der Schule wichtige Kooperation zwischen Lehrkräften gefördert werden kann.

Beiträge des Symposiums

Das technologiebezogene Fach- und fachdidaktisches Wissen von angehenden und aktiven Mathematiklehrkräften

Timo Kosiol, Stefan Ufer
Ludwig-Maximilians-Universität München

Das fachliche Professionswissen von Lehrkräften gilt als Voraussetzung für erfolgreiche Unterrichtsgestaltung. Shulman's (1986) Unterscheidung von fachlichem (CK), fachdidaktischem (PCK) und pädagogischem Wissen (PK) wurde von Koehler und Mishra (2009) um technologiebezogene Facetten erweitert: Das technologiebezogene Fachwissen (TCK) beispielsweise umfasst, wie Technologien genutzt werden können, um fachliche Inhalte darzustellen und fachlich zu arbeiten. Zum technologiebezogenen fachdidaktischen Wissen (TPCK) gehört, wie Technologien genutzt werden können, um Lerninhalte darzustellen, zu unterrichten und Lernprozesse zu unterstützen (Chai et al., 2013). Erprobte Testinstrumente für T(P)CK gibt es bisher keine und es wird üblicherweise mittels Selbsteinschätzungen mit generischem Fachbezug erhoben (Petko, 2020). In Studien zum Stand der Digitalisierung im Unterricht werden im Mittel günstige Selbsteinschätzungen für TPCK berichtet, fachspezifische digitale Werkzeuge werden jedoch selten im Unterricht verwendet (Drossel et al., 2019; Lorenz et al., 2021).

Die Entwicklung objektiver Testinstrumente für T(P)CK und die Untersuchung ihrer potenziellen Wirkung ist ein Forschungsdesiderat. Weiterhin fehlt Beschreibungswissen, über welches T(PCK) Lehrkräfte verfügen und welche Anforderungen sie in welcher Breite bewältigen.

Wir untersuchen ein neu entwickeltes Instrument zur Messung von mathematischem TCK und TPCK bezüglich (1a) seiner Reliabilität, (1b) der Authentizität und Relevanz der enthaltenen Anforderungen und beschreiben (2) welche Anforderungen Lehrkräfte zu diesen Wissensfacetten bewältigen können.

Das T(P)CK-Testinstrument fokussiert auf Tabellenkalkulationen, Dynamische Geometrie-Software (DGS) und Computer-Algebra-Systeme (CAS), da diese in einschlägigen Lehrplänen benannt sind (ISB, 2022). Die 26 TCK-Items enthalten technologische Anforderungen im Kontext der Schulmathematik, sind aber anders als die TPCK-Items nicht in unterrichtliche Kontexte eingebettet. Die 20 TPCK-Items umfassen technologiebezogene Anforderungen zu Instruktionsstrategien, Schülerkognitionen und Aufgabenpotential.

Die Tests wurden von N=205 Lehrkräften und N=189 Lehramtsstudierenden für weiterführende Schularten online bearbeitet. N=84 (TCK) bzw. N=74 (TPCK) der Lehrkräfte bewerteten die Tests hinsichtlich Authentizität und Relevanz für den Mathematikunterricht (1 „Trifft nicht zu“ bis 4 „Trifft zu“).

Offene Antworten (18 TCK- und 14 TPCK-Items) wurden mit zufriedenstellenden Interraterreliabilitäten kodiert. Die Ergebnisse einer Partial-Credit-Skalierung sprechen für einen

akzeptablen Itemfit. Eine Anforderung wurde ab einer vorhergesagten Lösungswahrscheinlichkeit von 62% als (mit hinreichender Sicherheit) beherrscht angesehen (vgl. OECD, 2021).

Die WLE-Reliabilitäten waren akzeptabel bis zufriedenstellend (1a, TCK: 0,66, TPCK: 0,72). Authentizität (M=3,11) und Relevanz (M=3,25) des TPCK-Test wurden sehr hoch eingeschätzt (1b), die Werte für den TCK-Test waren niedriger, aber im oberen Skalenbereich (Authentizität: M=2,36; Relevanz: M=2,61).

Für Fragestellung (2) wurden die Skalen jeweils in Quartile geteilt und analysiert, welche Items von einem Quartil bewältigt werden (d.h. Itemparameter < Personenparameter der unteren Quartilgrenze).

Bezüglich TCK können Dreiviertel der Teilnehmenden mindestens ein geeignetes digitales Werkzeug für eine schulmathematische Anforderung nennen. Während 50% der Teilnehmenden in einer DGS-Funktionen darstellen oder Schieberegler erzeugen können, wird das mehrschritte Konstruieren selbst vom obersten Quartil nicht sicher beherrscht. Es interpretiert jedoch einen CAS-Befehl angemessen und benennt zwei mit einem CAS bewältigbare mathematische Anforderungen. Die selbständige Nutzung einer Tabellenkalkulation wurde auch vom obersten Quartil nicht sicher beherrscht.

Bezüglich TPCK benennen 75% der Teilnehmenden mindestens einen Mehrwert einer Simulation, während 50% auch in verschiedenen Kontexten Mehrwerte digitaler Werkzeuge für bestimmte Lernziele angeben. Bei der fachdidaktischen Bewertung digitaler Aufgaben und Lernumgebungen gibt das oberste Quartil einen tragfähigen Aspekt an, eine differenzierte Bewertung von Stärken und Schwächen sowie die Konzeption eigener Lernumgebungen gelingt jedoch auch ihnen nicht sicher.

Das unseres Wissens nach erste Testinstrument für mathematisches TCK und TPCK hat sich als tragfähig erwiesen. Die vorliegende deskriptive Analyse liefert Einblicke, welche Anforderungen bezüglich der Wissensfacetten angehende und aktive Lehrkräfte bewältigen können. Obwohl dies in Lehrplänen gefordert wird (ISB, 2022), können viele Teilnehmenden nicht eigenständig mit fachlichen Technologien arbeiten. Das nur bei wenigen Teilnehmenden verfügbare Wissen zur fachdidaktischen Analyse digitaler Lernumgebungen ergänzt Befunde zur geringen Nutzungshäufigkeit (vgl. Drossel et al., 2019). Die Ergebnisse illustrieren, auf welchem Niveau Fortbildungsprogramme ansetzen könnten.

Die Rolle medienbezogener Lehrkompetenzen für die Medienkompetenz von Schülerinnen und Schülern

Karsten Stegmann[1], Tamara Kastorff[2], Frank Fischer[2]
[1]Universität Passau, [2]Ludwig-Maximilians-Universität München

Eine zentrale Forderung der Lehr-Lernforschung ist die Förderung medienbezogener Lehrkompetenzen von Lehrkräften. Die Förderung dieses Kompetenzbereiches wird als Schlüssel für die erfolgreiche Digitalisierung von Unterricht in der Schule angesehen. Bislang fehlen jedoch belastbare empirische Befunde, die die medienbezogene Lehrkompetenz von Lehrkräften mit der Qualität digital gestützten Unterrichts sowie den Kompetenzen der Lernenden in Beziehung setzen.

Von Lehrkräften mit höherer medienbezogener Lehrkompetenz kann z. B. erwartet werden, dass sie digitale Medien im Unterricht häufiger zur Förderung und Umsetzung konstruktiver und interaktiver Lernaktivitäten im Sinne des ICAP-Modells (Chi & Wylie, 2014) und seltener für passive und aktive

Lernaktivitäten einsetzen. Der Umgang mit digitalen Medien im Rahmen konstruktiver und interaktiver Lernaktivitäten sollte in einem stärkeren positiven Zusammenhang mit der Medienkompetenz der Schülerinnen und Schüler (SuS) stehen als passive und aktive Lernaktivitäten. Die medienbezogene Lehrkompetenz sollte daher vermittelt über Nutzung von Computern im Unterricht sowie die Qualität des Unterrichts auf die Medienkompetenz der SuS wirken. Daraus ergeben die folgenden Fragestellungen:

(RQ1) Inwieweit kann die medienbezogene Lehrkompetenz von Lehrkräften die Medienkompetenz von SuS vorhersagen?

(RQ2) Inwieweit kann der Umfang der Nutzung von Computern im Unterricht, die Qualität des Unterrichts und deren Interaktion die Medienkompetenz von SuS vorhersagen?

(RQ3) Inwieweit wird der ggf. vorhandene Zusammenhang (s. RQ1) durch den Umfang der Nutzung von Computern im Unterricht, die Qualität des Unterrichts und deren Interaktion mediiert?

Im Rahmen einer bayernweiten dreistündigen Datenerhebung im Herbst 2021 wurden von N = 2.421 SuS der 8. Jahrgangsstufe aus N = 133 Klassen an N = 39 Schulen (n = 13 Gymnasien, n = 13 Realschulen, 13 = Mittelschulen) sowie N = 181 Lehrkräfte (n = 72 Gymnasien, n = 49 Realschulen, 52 = Mittelschulen) Daten erhoben. Bei n = 1.470 SuS wurde die Medienkompetenz gemessen (NEPS ICT-Literacy; Senkbeil et al., 2013). Die Nutzung von Computern im Unterricht wurde mit einem Instrument aus der ICILS-Studie erfasst (vgl. Eickelmann et al., 2019). Die Qualität des Unterrichts wurde mit Hilfe eines für die Schule angepassten ICAP-Instruments erfasst (Stegmann, 2022). Die medienbezogene Lehrkompetenz von Lehrkräften wurde mit einem von Lachner und Kollegen (2019) entwickelten Test erfasst. Zur Beantwortung der Fragestellungen wurden hierarchische 3-Ebenen-Regressionsanalysen durchgeführt. Als Kontrollvariablen dienten die Schulart, der sozioökonomische Status (SES) sowie die private Verfügbarkeit von Computern.

Die Analysen bezüglich RQ1 ergaben keinen signifikanten Effekt der medienbezogenen Lehrkompetenz von Lehrkräften auf die Medienkompetenz der SuS. Damit entfällt die Mediatoranalyse (RQ3). Die Befunde bezüglich RQ2 zeigen negative Effekte der Nutzung von Computern im Unterricht sowie der Unterrichtsaktivitäten auf den ICAP-Leveln ‚passiv' und ‚aktiv'. Einen positiven Effekt auf die Medienkompetenz der SuS hatten hingegen Unterrichtsaktivitäten auf den ICAP-Leveln ‚konstruktiv' und ‚interaktiv'.

Die Befunde bezüglich RQ1 deuten darauf hin, dass die medienbezogenen Lehrkompetenz von Lehrkräften keine hinreichende Variable zur Erklärung der Medienkompetenz der SuS darstellt. Da ein großer Teil Varianz durch Faktoren außerhalb der Lehrkraft aufgeklärt wird (z. B. Schulart, SES, private Verfügbarkeit von Computern), ist dies auch nur bedingt verwunderlich. Umso wichtiger ist jedoch der Befund, dass die Qualität des Umgangs mit digitalen Medien im Unterricht dennoch Zusammenhänge mit der Medienkompetenz der SuS aufweist, da der Befund deutlich macht, dass Qualität der Aktivitäten im Unterricht und damit auch die medienbezogenen Lehrkompetenz von Lehrkräften im Zusammenhang mit der Medienkompetenz der SuS stehen. Es schließt sich daher die Frage an, inwieweit Kontextfaktoren wie die technische Ausstattung, die Einstellung der Schulleitung, etc. Gelingensbedingungen dafür darstellen, dass sich die medienbezogene Lehrkompetenz von

Lehrkräften tatsächlich in der Qualität des Umgangs mit digitalen Medien im Unterricht zeigt. Anstelle der aktuellen Medienkompetenz der SuS sollte in zukünftigen Analysen der Zuwachs der Medienkompetenz als Kriterium verwenden, da in querschnittlichen Daten die interindividuellen Unterschiede Effekt auf den Zuwachs leicht überdecken können.

Kooperation und Konzeption von virtuellen, schulübergreifenden Professionellen Lerngemeinschaften

Melanie Ripsam[1], Eike Klausing[2], Tobias Held[3], Claudia Nerdel[1], Cornelia Gräsel[2], Martin Lindner[3]
[1]Technische Universität München, [2]Bergische Universität Wuppertal, [3]Martin Luther Universität Halle-Wittenberg

Das Konzept der Professionellen Lerngemeinschaft (PLG) ist ein bekanntes Instrument der Schulentwicklung, steht jedoch durch die Möglichkeiten der (a)synchronen Zusammenarbeit (Stichwort: Distance Working) vor vollkommen neuen Chancen. Empirische Studien weisen auf positive Wirkungen hinsichtlich der Professionalisierung von Lehrkräften durch PLGs hin. Dabei kann eine PLG mittels 5-Komponenten-Modellen oder inhaltsanalytisch strukturierenden Ansatz beschrieben werden. Demnach wird zwischen den Dimensionen aktional, ideell und strukturell differenziert (Warwas & Schadt, 2020). Ferner werden mit Blick auf die konzeptionelle Gestaltung einer PLG die Ebenen Zustand und Prozess betrachtet. Soll die kooperative Entwicklung innerhalb der PLG fokussiert werden, so gewinnt nach Kansteiner und Kollegen (2020) die prozesshafte, aktionale Dimension an Bedeutung. Die virtuelle Zusammenarbeit per se bleibt bisweilen unberücksichtigt.

Spätestens seit Ausbruch der Corona-Pandemie sind Schulen gezwungen sich der technischen Revolution zu fügen. Vor allem Lehrkräfte stehen dabei immer wieder vor großen Herausforderungen. Daher müssen Schulen beim Einsatz digitaler Medien im MINT-Unterricht sowie der Weiterentwicklung ihrer Medienkonzepte unterstützt werden. Als wichtigen Erfolgsfaktor für die Schul- und Unterrichtsentwicklung sollte dabei die Zusammenarbeit der unterschiedlichen Akteure gestärkt werden. Daher untersucht DigitUS die förderlichen und hemmenden Bedingungen von Lerngemeinschaften für die Schul- und Unterrichtsentwicklung mit digitalen Medien. In diesem Zuge gilt es neben dem Zustand einer PLG mit seiner Struktur vor allem die Zusammenarbeit im virtuellen Setting (vPLG) zu betrachten. Es lassen sich folgende Fragestellungen ableiten:

1) Wie schätzen Teilnehmende der DigitUS-vPLG auf Basis ihrer eigenen Zielvorstellungen die konzeptionelle Umsetzung der vPLG ein?

2) Wie gestalten sich Kooperationsprozesse in der DigitUS-vPLG und inwieweit werden Entwicklungsprozesse auf den Dimensionen der Schulentwicklung angestoßen?

Zur Erfassung der notwendigen Voraussetzungen, Prozesse, Herausforderungen und Potentiale von derartigen Kooperationen wurden zwei schulübergreifende PLGs virtuell etabliert, in denen mit unterschiedlichen Expertisen an gemeinsam identifizierten Themen und Bedarfe gearbeitet wird: Eine davon erstreckt sich über zwei Bundesländer mit zwei Schulen aus Nordrhein-Westfalen und einer Schule aus Sachsen-Anhalt. Die andere setzt sich aus drei bayerischen Schulen zusammen. In den Lerngemeinschaften arbeiten schulische Teilnehmende, bestehend aus Vertreter*innen der Schulleitung, MINT-Lehrkräften und Wissenschaftler*innen (z.B. aus der Naturwissenschaftsdidaktik oder Schulentwicklung), zusammen. Des Weiteren nehmen Vertreter*innen aus der

Bildungsadministration und außeruniversitären Lehrerbildung teil. In sechswöchigen Abständen treffen sich die Akteur*innen der jeweiligen schulübergreifenden Lerngemeinschaften virtuell und planen ihr Vorgehen bei der (Weiter-)Entwicklung digitaler Unterrichts- und Schulkonzepte. Zwischen und während der Treffen kooperieren die Akteur*innen über eine virtuelle Plattform (Moodle).

Die ersten Lerngemeinschaften wurden von Mai 2021 bis August 2022 durchgeführt. Es wurden leitfadengestützte Interviews mit den Teilnehmenden geführt, wobei das erste Interview im Laufe der Kooperation in der PLG und ein zweites Interview im Anschluss der Kooperation stattfand. Es folgt die qualitative Auswertung der Interviewtranskripte nach Mayring (2010) mit MAXQDA.

Die Analysen deuten darauf hin, dass das Konzept von den Teilnehmenden als vielversprechend angesehen wird und die Lehrkräfte die konzeptionelle Umsetzung als virtuelle PLG positiv einschätzen. Mit Blick auf die Dimension Unterrichtsentwicklung scheinen bayerische Lehrkräfte verstärkt eine fachspezifische Zusammenarbeit (Fokus: Chemie) für die Entwicklung digitaler Unterrichtskonzepte zu wünschen. Die Ergebnisse liefern erste Hinweise, dass eine theoriegeleitete Vorgehensweise in der vPLG nötig ist. Dabei sollten jedoch vermehrt Praxisphasen zur Förderung von Kooperationsprozessen auf Unterrichtsebene (z.B. zur Gestaltung und Reflektion von Unterrichtsentwürfen mit verschiedenen digitalen Werkzeugen für das Erreichen fachlicher und prozessbezogener Lehr- und Lernziele) implementiert werden. Die virtuellen Treffen über Zoom wurden von den Teilnehmenden positiv bewertet, doch scheint die asynchrone Zusammenarbeit über Moodle ausbaufähig zu sein.

Basierend auf den Ergebnissen des Pilotprojekts wird die Konzeption der Lerngemeinschaft dahingehend optimiert, dass die Lehrkräfte vor allem effizient fachspezifisch zusammenarbeiten können. Ferner sollte die Kooperation über Moodle stärker angeleitet und mit weiterem digitalem Unterrichtsmaterial ausgebaut werden.

Die Rolle schulinterner Strukturen bei der Gründung professioneller Lerngemeinschaften zum erfolgreichen Einsatz von Bildungstechnologien - Eine qualitative Interviewstudie

Andrea Ludwig[1], Sonja Berger[2], Pauline Woop[2], Vera Haldenwang[1], Markus Teubner[1], Karsten Stegmann[3]
[1]Staatsinstitut für Schulqualität und Bildungsforschung München, [2]Ludwig-Maximilians-Universität München, [3]Universität Passau

Um dem gestiegenen Bedarf nach Kompetenzentwicklung von Lehrkräften beim Einsatz von Bildungstechnologien zu begegnen, sind Akteure der Lehrerbildung angehalten, Konzepte zu entwickeln, die es ermöglichen, „Lehrkräfte in fach- bzw. berufsfeldbezogenen Netzwerken (d.h. moderierten, fachbezogenen Communities of Practices) langfristig für fachbezogenes digital gestütztes Unterrichten [zu professionalisieren]“ (Ständige Wissenschaftliche Kommission, 2022). Professionelle Lerngemeinschaften (PLGen) von Lehrkräften, die, von Wissenschaftler:innen begleitet, über einen längeren Zeitraum hinweg zusammenarbeiten, sind eine Umsetzungsmöglichkeit dieser Communities. Unter Anwendung der grundlegenden Prinzipien einer PLG (Bonsen & Rolff 2006) soll der Kompetenzerwerb der Lehrkräfte in einer situierten Lernumgebung (vgl. Putnam & Borko, 2001) gestärkt werden. Gräsel et al (2006) haben in ihrer qualitativen Forschung zu PLGen bereits einige Barrieren für PLGen extrahiert. Eine jüngere Pilotstudie zur Etablierung von PLGen bekräftigt den höheren Aufwand personeller Ressourcen und betont neben der schulischen Ausstattung die Notwendigkeit der Unterstützung der Schulleitung zur Schaffung von schulinternen

Strukturen (Stegmann et al, 2021). Die vorliegende Interviewstudie im Rahmen des BMBF-geförderten Kooperationsprojektes DigitUS (Digitalisierung von Unterricht in der Schule) untersucht, welche schulinternen Strukturen zur Etablierung von PLGen zur Förderung des Einsatzes von Bildungstechnologien im Schulunterricht notwendig sind.

Dazu wurden 24 Lehrkräfte aus vier am Projekt teilnehmenden Sekundarschulen zu PLGen in halbstrukturierten Gruppeninterviews befragt. Die Stichprobe bestand aus fünf Multiplikator:innen, einer Schulleitung, einer Systembetreuung und 17 Lehrkräften, die mindestens eines der Fächer Mathematik, Biologie oder Natur und Technik unterrichteten und an einer PLG im Projektjahr 2021/2022 teilnahmen. Die Auswertung erfolgte anhand der qualitative Inhaltsanalysen nach Mayring (2015).

Folgende Gelingensbedingungen auf Ebene der Schule für die erfolgreiche Etablierung von PLGen konnten identifiziert werden.

Freiwilligkeit. In Einklang mit Gräsel et al. (2006) ist Freiwilligkeit eine grundlegende Bedingung für die erfolgreiche Kooperation von Lehrkräften in der PLG. So berichtete eine PLG, dass die von der Schulleitung angeordnete Teilnahme die Bereitschaft der Mitglieder, sich mit dem Thema digitale Medien im Schulunterricht auseinanderzusetzen und gemeinsam daran arbeiten zu möchten, stark einschränkte.

Materielle Ressourcen. Eine vielfältige technische Ausstattung bietet der PLG ein größeres Spektrum an Explorationsmöglichkeiten im Vergleich zu Schulen mit wenig technischer Ausstattung. Dies zeigt sich unter anderem daran, dass die Mitglieder einer befragten PLG die gute technische Ausstattung ihrer Schule als großen Vorteil beschrieben. An einer anderen Schule wurde über Schwierigkeiten berichtet, da die technischen Geräte (Tablets und Beamer) zwischen den Räumen transportiert werden musste und keine stabile Internetverbindung vorhanden war. Dies erhöhte den Aufwand und hinderte die PLG an ihren Aktivitäten.

Räumliche Ressourcen. Für die PLG reservierbare Räume stellten sich als bedeutsames Kriterium heraus, da die PLG einen Rückzugsort außerhalb des Lehrerzimmers benötigte, um einen sicheren Raum zu schaffen. Beispielsweise berichtete eine PLG, dass die Klassenräume der teilnehmenden Lehrkräfte abwechselnd für Treffen zur Verfügung gestellt wurden, wodurch stets gewährleistet war, dass ausreichend Platz und technische Mittel für Treffen zur Verfügung standen.

Zeitliche Ressourcen. Lehrkräfte betonten, dass die Unterstützung durch die Schulleitung entscheidend sei, um den erhöhten Zeitaufwand durch Entlastung an anderer Stelle zu kompensieren. Gemeinsame Freistunden am Vormittag, während dieser die PLG tagen kann, erfordern von den Teilnehmenden keine Auslagerung der PLG-Aktivitäten in die Freizeit. Die Gewährung von Anrechnungsstunden wirke sich zudem positiv auf die Bereitschaft zur Teilnahme an den Aktivitäten der PLG aus.

Die Ergebnisse der Interviewbefragung deuten darauf hin, dass zur erfolgreichen Etablierung einer PLG an der Schule neben der Freiwilligkeit der Teilnahme insbesondere Ressourcen in den zeitlichen, räumlichen und materiellen Rahmenbedingungen vorhanden sein müssen. Damit werden die bereits bekannten Gelingensbedingungen (Gräsel et al, 2006; Stegmann et al, 2021) um

konkretere Umsetzungsmöglichkeiten zur Verbesserung der schulinternen Strukturen ergänzt. Aus diesen lassen sich als Entscheidungsgrundlage geeignete Handlungsempfehlungen für Akteure der Bildungsadministration, die den Einsatz von PLGen erwägen, ableiten.

A-20: Zusammenhänge zwischen der Kompetenz von Lehrkräften, Selbstregulation zu fördern, und der Selbstregulation ihrer Lernenden

Zeit: Dienstag, 28.02.2023: 13:00 - 14:45 · *Ort:* R11 T00 D05

Zusammenhänge zwischen der Kompetenz von Lehrkräften, Selbstregulation zu fördern, und der Selbstregulation ihrer Lernenden

Chair(s): **Charlotte Dignath** (TU Dortmund), **Yves Karlen** (Pädagogische Hochschule FHNW, Switzerland)
*Diskutant*in(nen):* **Joachim Wirth** (Ruhr-Universität Bochum)

Das Forschungsfeld zu selbstreguliertem Lernen (SRL) hat sich in den vergangenen 30 Jahren deutlich erweitert und verändert. Während die frühe Forschung zu SRL in den 1980er bis 2000er Jahren noch sehr stark auf die Lernenden fokussierte, interessierte sich die jüngere Forschung ab 2000 zunächst nur langsam auch für die Lehrkräfte. Erst im letzten Jahrzehnt sind die Lehrkräfte und ihre Förderung von SRL in den Vordergrund von Studien gerückt. Nur sehr selten allerdings wird die Forschung zum SRL von Lernenden mit der Forschung zur SRL-Förderung durch Lehrkräfte zusammengebracht.

Im vorliegenden Symposium sollen diese beiden Aspekte – die Förderung von SRL durch Lehrkräfte und die SRL der Lernenden – kombiniert werden, um Wirkzusammenhänge zwischen dem Lehrkräfteverhalten und dem Verhalten der Schüler*innen identifizieren zu können. Dabei kommen in den vier interdisziplinären Beiträgen aus der Entwicklungspsychologie, der Pädagogischen Psychologie und der Erziehungswissenschaft verschiedene Forschungsmethoden und Designs (z. B. Strukturgleichungsmodelle, Experience Sampling, Literaturreview) zum Einsatz.

In der ersten Studie des Symposiums werden Zusammenhänge zwischen der beruflichen Selbstregulation von Lehrkräften und dem selbstberichteten selbstregulierten Lernen ihrer Schüler*innen vorgestellt. In der zweiten Studie wird geprüft, wie die beiden Kompetenzbereiche der Lehrpersonen als selbstregulierte Lernende einerseits und als selbstregulationsfördernde Lehrkraft andererseits mit ihrer selbstberichteten SRL-Förderung im Unterricht und auch den von den Schüler*innen selbstberichteten SRL-Kompetenzen zusammenhängt. Als drittes wird eine Experience Sampling-Studie präsentiert, in der die Selbstregulation von Schüler*innen mithilfe eines Fragebogens als Trait-Maß, sowie mithilfe von Experience Sampling als State-Maß erfasst wird. Neben Zusammenhängen zwischen diesen beiden Maßen ermitteln sie zudem, wie die Selbstregulation der Schüler*innen mit der von ihnen wahrgenommenen Unterrichtsqualität zusammenhängen. Als letztes wird ein Literaturreview vorgestellt, der das Ziel hat, die Forschung der vergangenen zwanzig Jahre zur Förderung von SRL durch Lehrkräfte anhand des Coactiv-Modells zu strukturieren, um einen Überblick über die Evidenz zu Determinanten und Konsequenzen der professionellen Kompetenz von Lehrkräften bzgl. der SRL-Förderung zu geben und Forschungsbedarfe abzuleiten.

Gesamthaft heben die Beiträge die Bedeutung der professionellen Kompetenzen der Lehrpersonen im SRL für das SRL der Lernenden hervor. Zugleich wird deutlich, dass es weitere Forschung braucht, um das Zusammenspiel verschiedener Kompetenzaspekte bei Lehrkräften und zwischen Lehrkräften und Lernenden besser zu verstehen, um daraus Implikationen für die professionelle Entwicklung von Lehrkräften im SRL abzuleiten.

Beiträge des Symposiums

Wie modellhaft können Klassenlehrkräfte für das selbstregulierte Lernen von Viert- und Fünftklässler*innen sein?

Manuela Benick, Franziska Perels
Universität des Saarlandes

Entsprechend der sozial-kognitiven Lerntheorie Banduras (1986) entwickelt sich die Fähigkeit zur Selbstregulation entlang der vier hierarchischen Niveaustufen Beobachtung, Nachahmung, Selbstkontrolle und Selbstregulation und wird maßgeblich durch die Beobachtung und Nachahmung sozialer Modelle (z. B. Lehrkräfte oder Peers) beeinflusst (Schunk & Usher, 2018). In der aktuellen Diskussion über wirksame Maßnahmen zur Förderung selbstregulierten Lernens (SRL) bei Schüler*innen wird Lehrkräften folglich eine Schlüsselfunktion zugeschrieben, da sie über die Vermittlung und Förderung von SRL-Strategien entscheiden (Donker et al., 2014;) und als Modelle für ihre Schüler*innen fungieren (Bembenutty, 2013; Moos & Ringdal, 2012). Über Art und Häufigkeit der Instruktion kognitiver SRL-Strategien durch Lehrkräfte während des Unterrichts liegen bereits Befunde vor (Dignath & Büttner, 2018; Kistner et al., 2010) ebenso wie zu den Einstellungen von Lehrkräften zur Bedeutsamkeit von SRL für das schulische Lernen (Dignath & van der Werf, 2012; Ruprecht et al., 2017). Doch wie sieht es mit den selbstregulativen Fähigkeiten der Lehrkräfte selbst aus? Sind Lehrkräfte erfolgreich Selbstregulierende und können sie daher auch modellhaft für ihre Schüler*innen sein?

Um erste Anhaltspunkte zur Beantwortung dieser Fragen zu eruieren, wurden in einem querschnittlichen Design die Daten von N = 43 Klassenlehrkräften (83.70% weiblich; MAlter = 40.42 Jahre, SD = 9.86, Range = 27-63 Jahre; n = 30 Lehrkräfte Elementarbereich (vierte Klassenstufe), n = 13 Lehrkräfte Sekundarbereich (fünfte Klassenstufe)) und deren Schüler*innen mehrebenenanalytisch betrachtet.

Basierend auf dem Selbstregulationsansatz von Zimmerman (2000) lag der Fokus dabei auf metakognitiven und motivationale Komponenten der schulischen bzw. beruflichen Selbstregulation von Schüler*innen bzw. Lehrkräften. Der Begriff der beruflichen Selbstregulation zielt dabei auf die Arbeitswelt von Lehrkräften ab und betrachtet Selbstregulation im Kontext von Tätigkeiten, mit denen Lehrkräfte in ihrem Berufsalltag konfrontiert sind wie Unterrichtsvorbereitungen, Unterrichtsdurchführung und -nachbereitung, Korrekturen, Materialrecherche und -erstellung. Das selbstregulierte Lernverhalten der Schüler*innen wurde gleichermaßen erfasst, indem n = 440 Viertklässler*innen (49.30% weiblich; MAlter = 9.84 Jahre, SD = 0.57, Range = 9-12 Jahre) und n = 255 Fünftklässler*innen (44.90% weiblich; MAlter = 10.81 Jahre, SD = 0.54, Range = 10-13 Jahre) zu metakognitiven und motivationalen Aspekten ihres schulischen Lernens via Selbstbericht befragt wurden. Der Fragebogen für Schüler*innen (Benick et al., 2021) umfasste N = 45 Items zur Erfassung der Variablen Zielsetzung, Zeitplanung, Strategieplanung, Selbstwirksamkeit, Selbstmotivation, Aufmerksamkeitsfokussierung, Selbstbeobachtung, Selbstevaluation und Kausalattribution (Alpha = 0.73 – 0.92). Der Fragebogen zur beruflichen Selbstregulation von Lehrkräften zielte auf die Erfassung selbiger Variablen ab und umfasste N = 44 Items (Alpha = 0.79 – 0.88). Zusätzlich zu diesem Selbstberichtsverfahren wurden auf Ebene der Lehrkräfte Fallvignetten verwendet, um ihr Strategiewissen hinsichtlich SRL zu erfassen (Cleary & Callan, 2018). Erste Bestimmungen der Intraklassenkorrelationskoeffizienten für die jeweiligen SRL-Variablen deuten darauf hin, dass die

berufliche Selbstregulation von Lehrkräften einen Einfluss auf das selbstregulierte Lernen von Schüler*innen ausübt (ICC > 0.05; Muthén & Satorra, 1995).

Practise what you preach: Selbstreguliertes Lernen der Lehrpersonen und deren Zusammenhänge zu den Kompetenzen der Schüler:innen

Yves Karlen[1], Silke Hertel[2], Johannes Jud[1], Carmen Hirt[1]
[1]Pädagogische Hochschule FHNW, [2]Universität Heidelberg

Die bisherige Forschung betont die Bedeutung von Lehrpersonen bei der Förderung des selbstregulierten Lernens (SRL) von Schüler:innen (Dignath & Veemann, 2021) und hebt die Notwendigkeit derer professionellen Handlungskompetenzen in diesem Bereich hervor (Kramarski & Hearysman, 2021). Dabei wird insbesondere zwischen den Erfahrungen der Lehrperson als selbstreguliert Lernende sowie ihren Kompetenzen in der konkreten Vermittlung und Förderung von SRL unterschieden (Dembo, 2001; Karlen et al., 2020). Erstere beziehen sich auf Erfahrungen hinsichtlich der Nützlichkeit von SRL sowie der Herausforderungen beim SRL während der eigenen Lernprozesse. Letztere beihnalten Überzeugungen, Motivation und Wissen in Bezug auf die SRL-förderliche Unterrichtsgestaltung und die direkte Vermittlung von SRL im Regelunterricht. Die bisherige Forschung hebt die Bedeutung beider Kompetenzbereiche für das Handeln der Lehrpersonen hervor (Karlen et al., 2020). Dabei wird insbesondere auf die Relevanz der Überzeugungen und der Selbstwirksamkeit für die Erklärung von Unterschieden bei der Förderung von SRL im Unterricht hingewiesen (z. B. Lawson et al., 2019). Die positiven Effekte eines SRL-förderlichen Unterrichts für das SRL und den Lernerfolg der Schüler:innen wurde bereits mehrfach empirisch belegt (z. B. de Boer et al., 2018). Allerdings wurden die SRL-Kompetenzen der Lehrpersonen und der Schüler:innen bisher vorwiegend getrennt untersucht. Folglich ist es ein Desiderat der SRL-Forschung, das Zusammenspiel der professionellen Handlungskompetenzen von Lehrpersonen und der SRL-Kompetenzen der Schüler:innen differenziert zu analysieren.

Basierend auf einem Mehrebenenansatz untersucht diese Studie (1) das Zusammenspiel von professionellen Handlungskompetenzen der Lehrpersonen mit ihrem SRL-Handeln im Unterricht und (2) über dieses vermittelt die Bedeutung für die SRL-Kompetenzen der Schüler:innen.

Im Rahmen einer querschnittlich angelegten Studie wurden N = 167 Lehrpersonen (44.4% weiblich; MAlter = 44.34 Jahre, SD = 10.82) mit einer durchschnittlichen Berufserfahrung von 16.66 Jahren (SD = 10.92) aus 17 Schule der Sekundarstufe I befragt. Zudem nahmen deren N = 2 757 Schüler:innen (48.9% weiblich; MAlter = 14.45 Jahre, SD = 0.96) an der Befragung teil.

Mit Online-Fragebogenerhebungen und Wissenstests wurde das eigene SRL der Lehrpersonen (vier Subskalen: prozedurales metakognitives Wissen sowie metakognitive, kognitive und motivationale Regulationskompetenzen) und deren Kompetenzen als Vermittler:innen von SRL (PCK-SRL, Selbstwirksamkeit und Überzeugungen zur Relevanz von SRL) reliabel erfasst (Cronbach`s Alpha = .80 - 85.). Das Unterrichtshandeln der Lehrperson wurde mit einer Skala zu ihren SRL-Fördertätigkeiten erfasst (Cronbach`s Alpha = .72). Bei den Schüler:innen wurden das metakognitive Strategiewissen, die Nutzung von metakognitiven Strategien sowie kognitive und motivationale Regulationskompetenzen (Cronbach`s Alpha = .69 - .85.) reliabel erhoben. Als Kovariaten wurden das Geschlecht, die Anzahl Bücher und die Kulturgüter der Schüler:innen erhoben.

Die mehrebenenanalytische Modellierung (Chi2 = 111.561, df = 64, RMSEA = .016, CFI = .994; SRMRwithin = .008; SRMRbewteen = .094) ergab signifikante positive Effekte der Überzeugungen und der Selbstwirksamkeit der Lehrperson auf ihre SRL-Förderung. Zudem zeigte sich ein indirekter Effekt der eigenen SRL-Kompetenzen der Lehrpersonen vermittelt über die Selbstwirksamkeit auf ihre SRL-Förderung. Des Weiteren ergaben sich positive Korrelationen zwischen der SRL-Förderung der Lehrpersonen und den SRL-Kompetenzen der Schüler:innen (β = .327 Nutzung von Planungsstrategien; (β = .249 Nutzung von Monitoringstrategien; (β = .378 Nutzung von Evaluationsstrategien; (β = .320 motivationale Regulationskompetenzen).

Diese Ergebnisse unterstreichen die Bedeutung der beiden Kompetenzbereiche der Lehrpersonen als selbstregulierte Lernende sowie als Vermittler:innen von SRL für ihre SRL-Förderung im Unterricht. Zudem heben sie die Relevanz der SRL-Förderung im Unterricht für den Aufbau von SRL-Kompetenzen der Schüler:innen hervor. Die Ergebnisse machen deutlich, dass zukünftige Weiterbildungen für Lehrpersonen beide Komptenzbereiche vermitteln sollten.

Supporting individual students' daily self-regulation: Within-person level associations with instructional quality

Friederike Blume, Florian Schmiedek
DIPF | Leibniz Institute for Research and Information in Education

Individuals engaging in self-regulation (SR) control, direct, and plan cognitions, emotions, and behaviour (e.g., McClelland & Cameron, 2011). Studies considering SR as a trait largely stable over time emphasised the relevance of between-person differences to explain differences in academic outcomes (e.g., Duckworth et al., 2019). Recently, SR was also found to vary within students from day to day and associations with academic success were shown to be evident at the within-person level (e.g., Blume et al., 2022).

Aspects of instructional quality (IQ), and particularly teacher support (TS) and classroom management (CM), but not cognitive activation (CA), have been found to be positively associated with students' trait SR, suggesting their potential to support its development (e.g., Sankalaite et al., 2021). However, associations between students' subjectively perceived IQ and SR have not been examined. Also, whether associations between IQ and students' SR are evident at the within-person level and with respect to individual lessons is yet to be investigated. Finally, it is yet to be determined whether the strength of these associations is influenced by the level of students' trait SR.

Methods

64 students (Mage = 11.71 years, SD = 0.93) reported about their trait SR (Blume et al., 2020) and demographics in a background questionnaire. 57 of them participated in a 15-day daily diary phase in which students informed each afternoon on their perceived IQ (Talić et al., 2022) and their SR (adapted from Blume et al., 2020) in the last lesson that day over 15 consecutive school days. Multilevel models using Bayesian estimators and with students' SR in individual lessons as the outcome variable were estimated. Information with respect to the lesson level were centred on person means and individual mean scores over the study period were centred on grand means.

Results

Results indicated students' trait SR to be significantly associated with their SR in individual lessons (β = 0.54, 95% CI [0.32, 0.72]). Additionally, perceived TS (β = 0.18, 95% CI [0.05, 0.30]) and CM (β = 0.19, 95% CI [0.05, 0.32]), but not CA (β = -0.01, 95% CI [-0.06, 0.25]), were significantly associated with their SR at the lesson level. None of the interactions between lesson-related IQ and trait SR significantly predicted students' SR at the lesson level. Predictors accounted for 55.60% of the Level 1 variance in SR. Patterns of associations were largely the same for less complex models including trait SR and predictors related to one IQ dimension only and in relation to Level 2.

Discussion

In line with earlier evidence, findings indicated students' trait SR to be positively associated with their SR at the lesson level (e.g., Blume et al., 2021), emphasising the need for continuous support of students' with below-average SR in class.

Furthermore in line with previous observational studies considering the between-person perspective (cf., Sankalaite et al., 2021), the present findings showed that lessons with higher perceived TS and CM, but not CA, were lessons with higher SR. Although future studies need to establish causality, these findings might emphasise the relevance of perceived IQ for SR at the lesson level and suggest that more consideration should be given to students' perspectives. Additionally, they may be considered as indicative of the relevance of TS and CM to students' SR in individual lessons and thus of their potential for continuous support of individual student functioning in class.

Finally, findings showed students' trait SR to be largely irrelevant for the strength of associations between IQ and SR at the lesson level. They thereby may be considered to indicate that high IQ should be considered equally important for the SR of all students.

Anwendung des Coactiv-Modells zur Erklärung professioneller Kompetenz von Lehrpersonen hinsichtlich der Förderung von Selbstregulation beim Lernen – ein Review

Charlotte Dignath[1], Mareike Kunter[2]
[1]TU Dortmund, [2]DIPF | Leibniz Institute for Research and Information in Education

Theorie

Selbstregulation des Lernens (SRL) ist eine zentrale Fähigkeit für lebenslanges Lernen. Allerdings zeigt sich, dass viele Schüler*innen Selbstregulationsstrategien nur unzureichend beherrschen. Diese Unterschiedlichkeit in der Selbstregulation ist häufig auf unzureichendes Strategiewissen und mangelnde Erfahrung mit SRL zurückzuführen. Dabei zeigt sich auch, dass SRL erlernbar ist und mithilfe verschiedener Methoden gefördert werden kann. Das Fördern von SRL sollte daher ein wichtiges Ziel von Unterricht sein.

Während sich frühere SRL-Forschung vor allem mit der Frage der Wirksamkeit und Trainierbarkeit von SRL beschäftigt hat, gibt es erst in der jüngeren Zeit Forschung zur Rolle von Lehrkräfte bei der Entwicklung und Förderung der SRL von Schüler*innen. Um einen umfassenden Überblick über bisherige Forschung zu diesen Fragen zu erhalten und Forschungslücken zu identifizieren, wurde ein Literaturreview durchgeführt, der die vorliegende Evidenz anhand des COACTIV-Modells einordnet. Die Forschungsfragen der einbezogenen Studien untersuchten, wie gut Lehrpersonen in der Lage sind, SRL im Unterricht zu fördern und Erklärungsansätze für Unterschiede zwischen Lehrpersonen

zu entwickeln. Ziel des Reviews ist es zu verstehen, wie man die SRL von Schüler*innen fördern kann und Implikationen für weitere Forschung und Praxis abzuleiten.

Methode

Das Coactiv-Modell spezifiziert Determinanten und Konsequenzen professioneller Kompetenz von Lehrpersonen. Determinanten beinhalten dabei sowohl individuelle als auch Kontextbedingungen, und die Bereitstellung von Lerngelegenheiten für Lehrpersonen sowie deren Nutzung dieser Lerngelegenheiten, und Effekte auf die Entwicklung ihrer Kompetenz. Nach dem Coactiv-Modell gibt es verschiedene Aspekte professioneller Kompetenz, mit deren Hilfe das Lehrverhalten im Unterricht vorhergesagt werden kann: professionelles Wissen, Überzeugungen, motivationale Orientierungen und die Selbstregulation. Neben dem Unterrichtsverhalten stellen auch Outcome-Variablen auf Ebene der Schüler*innen sowie auf Ebene der Lehrpersonen langfristige Konsequenzen der professionellen Kompetenz dar (Kunter et al., 2011). Entsprechend dem Coactiv-Modell wurden die hier identifizierten Studien sortiert nach den Forschungsfragen, (1) welche Aspekte professioneller Kompetenz das Unterrichtsverhalten der Lehrpersonen mit Blick auf die Förderung von SRL vorhersagen, (2) wie das Unterrichtsverhalten bzgl. SRL mit verschiedenen Variablen der Schüler*innen und der Lehrpersonen zusammenhängen, (3) wie individuelle und kontextuelle Bedingungsfaktoren mit den Kompetenzaspekten zusammenhängen, und (4) welche Lerngelegenheiten effektiv sind für die Entwicklung professioneller Kompetenz bzgl. SRL und wie Lehrpersonen diese Lerngelegenheiten nutzen.

Ergebnisse und Diskussion

(1) Die Ergebnisse zeigen, dass die Überzeugungen von Lehrkräften in der Regel nur mit ihrer selbstberichteten SRL-Förderung aber nicht mit einer aus Unterrichtsbeobachtungen hergeleiteten SRL-Förderung zusammenhängen, und dass Lehrpersonen zwar meist positive Überzeugungen bzgl. SRL haben, aber tendenziell eine niedrige Selbstwirksamkeitserwartung aufweisen, SRL in ihrem Unterricht effektiv fördern zu können. Nur wenige Forschende haben bisher die eigene Selbstregulation der Lehrpersonen mitberücksichtigt. Während die Überzeugungen und motivationalen Orientierungen der Lehrpersonen traditionell mit Fragebogen erfasst wurden, bildet eine valide Erfassung der Selbstregulation der Lehrpersonen eine Herausforderung. Auch die Erfassung des professionellen Wissens von Lehrpersonen zu SRL und zur Förderung von SRL stellt aktuell noch ein messmethodisches Problem dar.

(2) Während es mittlerweile einige Studien gibt, in denen die SRL-Förderung von Lehrpersonen im Unterricht multimethodal erfasst wurde, werden nur selten Outcomevariablen auf Ebene der Schüler*innen hinzugezogen, wodurch dieser Forschungsbereich bisher v.a. querschnittlich und damit eher deskriptiv bleibt.

(3) Hinsichtlich der Kontextbedingungen wurde gezeigt, dass neben den individuellen Überzeugungen zu SRL auch die gemeinsame SRL school vision eine Rolle spielt bei der Entwicklung professioneller Kompetenz bzgl. SRL. Daneben hat die SRL-Forschung zu Lehrpersonen die Kontextbedingungen bisher nur wenig berücksichtigt.

(4) Die Trainingsstudien, in denen die Wirksamkeit von Lerngelegenheiten zum Thema SRL für Lehrpersonen untersucht wurden, fokussieren meist auf Lehramtsstudierende und beziehen nur selten Lehrkräfte im Schuldienst mit ein.

Die Ergebnisse bieten einen spezifischen Überblick über Forschung zu Determinanten und Konsequenzen professioneller Kompetenz von Lehrpersonen hinsichtlich der Förderung von SRL, und liefern konkrete Einblicke in Forschungslücken, die es noch zu füllen gilt.

Session B

Einzelbeiträge

B-01: Kooperation mit Eltern

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S06 S00 B29
Chair der Sitzung: Kristina Ackel-Eisnach

Emotionen von Eltern in schulischen Lehrer-Eltern-Gesprächen

Kristina Ackel-Eisnach, Frank Behr, Josef Strasser
Universität Koblenz - Landau, Deutschland

Aus Sicht der Eltern sind Gespräche, die sich auf ihr eigenes Kind konzentrieren, immer emotional bedeutsam (Esquivel et al., 2008). Daher ist es nicht verwunderlich, dass Gespräche und Interaktionen zwischen Lehrpersonen und Eltern teils als emotional sehr belastend erlebt werden (Lasky, 2000). Im Hinblick auf eine gelingende Gesprächsführung ist folglich eine zentrale Anforderung an Lehrpersonen, auch in belastenden Gesprächen die Perspektive der Eltern und ihre Emotionen adäquat zu beachten (Aich & Behr, 2015). Da (antizipierte) negative Emotionen und damit verbundener Stress empathische Fähigkeiten erheblich beeinträchtigen (Tomova et al., 2014), ist es für Lehrkräfte wichtig zu wissen, mit welchen Emotionen sie in Elterngesprächen zu rechnen haben. Diesbezügliche realistische Erwartungen sind Voraussetzung der geforderten internalen und externalen Emotionsregulation (Göhringer et al., 2020). Es bedarf demnach gesicherten Wissens zur Perspektive und den Emotionen der Eltern in diesen Gesprächen.

Hier knüpft der Beitrag mit folgenden Fragestellungen an:

1. Mit welchen Gefühlen gehen Eltern in verschiedenen Gesprächssituationen in Lehrer-Eltern-Gespräche?
2. Besteht ein Zusammenhang zwischen den elterlichen Emotionen und ihren Haltungen und Einstellungen zu Schule und Lernen?
3. Wie beeinflussen die Emotionen der Eltern die Gesprächsführung und -gestaltung seitens der Lehrkräfte?

Es wurden Daten einer Elternbefragung aus dem Projekt „Lehrer-Eltern-Gespräche aus Elternperspektive (LEGEP)“ analysiert. Die Befragung wurde als Online-Querschnittserhebung mit Eltern schulpflichtiger Kinder verschiedener Schulformen durchgeführt. Insgesamt haben 1876 Eltern den Fragebogen bearbeitet. 27 % der Angaben können Elterngesprächen in der Primarstufe, 73 % Gesprächen in der Sekundarstufe zugeordnet werden. Die Emotionen der Eltern wurden über eine offene Frage erfasst und mithilfe einer inhaltlich strukturierenden Inhaltsanalyse nach Mayring (2015) kodiert. Auf einer ersten Ebene erfolgte eine dimensionale Unterscheidung nach drei Hauptkategorien (negativ, gemischt, positiv) und auf einer zweiten Ebene nach konkreten Emotionen (Brandstätter et al., 2018; Izard, 1999). Darüber hinaus umfasste das Befragungsinstrument Items zu Schüler- und Elternvariablen, zu Einstellungen und Haltungen der Eltern, zu Erwartungen und Wünschen, zu Gesprächsthemen und -zielen sowie der Zufriedenheit mit dem Gespräch. Cronbachs α lag für alle gebildeten Skalen in einem akzeptablen bis exzellenten Bereich.

Insgesamt benannten 619 Eltern Emotionen, die inhaltsanalytisch kategorisiert werden konnten. 59 % gaben positive (Freude und Interesse), 28 % negative (Furcht, Kummer, Zorn) und 13 % gemischte Emotionen an. Schulformbezogene Unterschiede waren nicht vorhanden. Eltern, die die Zusammenarbeit von Schule und Elternhaus positiver bewerteten, gingen häufiger mit positiven Emotionen in die Gespräche ($b = .38$, Waldχ^2 (1) = 23.73, $p < .05$), ebenso Eltern, die eine hohe schulische Selbstwirksamkeit aufwiesen ($b = .30$, Waldχ^2 (1) = 6.02, $p < .05$). Zudem zeigte sich, dass

Rückmeldegespräche häufiger mit positiven, Problem- und Konfliktgespräche jedoch eher mit negativen Emotionen einhergingen. Die Ergebnisse weisen weiter daraufhin, dass bei positiven Emotionen zu Beginn der Gespräche, die Zufriedenheit am Ende der Gespräche am höchsten war ($F(2,216) = 27.77$, $p < 0.05$, $eta^2 = .08$). Allerdings fällt auf, dass knapp 40 % der Eltern, die negative Emotionen angaben, am Ende der Gespräche eine hohe Zufriedenheit aufwiesen, wohingegen 11 % der Eltern, die mit positiven Gefühlen in die Gespräche gingen, unzufrieden mit den Gesprächsverläufen waren. Hier soll noch genauer untersucht werden, ob dies in Zusammenhang mit dem Lehrerhandeln in den Gesprächen oder den Erwartungen und Wünschen der Eltern steht.

Die Ergebnisse geben Hinweise darauf, in welchen Situationen und unter welchen Umständen Lehrkräfte mit eher positiv oder negativ gefärbten Emotionen rechnen können. Daraus ergeben sich Ansatzpunkte für eine adäquate Vorbereitung der Lehrkräfte sowie für die Schaffung günstiger Bedingungen auf der schulischen Ebene (etwa durch die Stärkung der Kooperation zwischen Schule und Elternhaus).

Hilfe – unser Kind hinkt hinterher!? Der Zusammenhang zwischen elterlicher Wahrnehmung des Sprachförderbedarfs, ihren Home-Literacy Aktivitäten, und den sprachlichen Fähigkeiten der Kinder

Jan Lenhart[1], Klaus Lingel[2]
[1]Otto-Friedrich-Universität Bamberg, Germany; [2]Julius-Maximilians-Universität Würzburg, Germany

Theoretischer Hintergrund und Fragestellung. Die familiäre Lernumwelt gilt als zentral für die kindliche Entwicklung und ist ein wichtiger Prädiktor des späteren schulischen Erfolges (z.B. Leseman & de Jong, 1998; Sénéchal, 2006; Storch & Whitehurst, 2002). Während sich das Wissen über den Beitrag spezifischer Home-Literacy Aktivitäten in den letzten Jahren zunehmend ausdifferenziert hat (z.B. Krijnen et al., 2020), gibt es vergleichsweise wenig Forschung zum Einfluss elterlicher Motive, Ziele und Überzeugungen auf die Gestaltung der häuslichen Lernumwelt, was insbesondere für das kompensatorische Fördermotiv (d.h. zum Ausgleich wahrgenommener sprachlicher Defizite bei den Kindern) gilt. Dabei ist es aus praktischer Perspektive hochrelevant, ob Eltern sprachlichen Förderbedarf bei ihren Kindern erkennen und wie sich ihre Wahrnehmung des Förderbedarfs auf die Gestaltung der familiären Lernumwelt, insbesondere konkrete Home-Literacy Aktivitäten, auswirkt.

Methode. Um einen ersten Einblick in dieses Thema zu erhalten, wurden 193 Kindergartenkinder hinsichtlich ihrer sprachlichen Kompetenzen (rezeptiver und expressiver Wortschatz, grammatikalische Kompetenzen, phonologische Bewusstheit, phonologisches Arbeitsgedächtnis und Erzählkompetenzen) untersucht. Ihre Eltern wurden in einem umfangreichen Fragebogen zur familiären Lernumwelt (kulturelle Ressourcen, eigenes Leseverhalten, informelle, sprachliche und schriftsprachliche Home-Literacy-Aktivitäten, Lob für Sprachfortschritte und Ermutigung zu Aktivitäten) und ihren Fördermotiven im Bereich der sprachlichen Entwicklung befragt. Anhand des Elternberichts über sprachlichen Förderbedarf wurden die Kinder in zwei Gruppen eingeteilt (ohne vs. mit berichtetem Förderbedarf).

Ergebnisse. Die Ergebnisse der Studien zeigen, dass sich die elterliche Wahrnehmung von sprachlichem Förderbedarf in den sprachlichen Kompetenzen ihrer Kinder widerspiegelten. Kinder von Eltern, die sprachlichen Förderbedarf berichteten, erzielten im mittleren Effektstärkebereich

signifikant niedrigere Werte hinsichtlich des expressiven und rezeptiven Wortschatzes, der grammatikalischen Kompetenz, der phonologischen Bewusstheit, des phonologischen Arbeitsgedächtnisses und der Erzählkompetenz (d = 0.45 bis 0.68). Zudem berichteten Eltern, die sprachlichen Förderbedarf angaben, häufiger für Fortschritte im Sprachbereich zu loben und zu mehr (schrift-)sprachlichen Aktivitäten zu ermuntern (d = 0.34 und 0.50). Interessanterweise zeigten sich jedoch keine signifikanten Unterschiede in den berichteten kulturellen Ressourcen, dem elterlichen Leseverhalten sowie den informellen, den sprachlichen und den schriftsprachlichen Home-Literacy Aktivitäten (d = -0.19 bis 0.19). In der Gruppe mit berichtetem Förderbedarf zeigten sich zudem in der Regel niedrigere und teilweise sogar negative Korrelationen zwischen Home-Literacy Aktivitäten und den sprachlichen Kompetenzen. Die Ergebnisse der Studie liefern somit erste Hinweise, dass Eltern sprachlichen Förderbedarf bei ihren Kindern erkennen und möglicherweise mit einer Steigerung der Home-Literacy Aktivitäten reagieren. Obwohl diese Interpretation inhaltlich plausibel scheint, ist aufgrund des querschnittlichen, korrelativen Designs ein Nachweis der Wirkrichtung jedoch nicht möglich und Bedarf weiterführender längsschnittlicher Forschung zur Untersuchung der Wechselwirkung zwischen dem Erkennen von sprachlichem Förderbedarf, der Anpassung von Home-Literacy Aktivitäten und der kindlichen Sprachentwicklung.

Die Auswirkungen des ersten Corona-Lockdowns auf die Mathematikkompetenzen von Grundschülerinnen und -schülern in Deutschland – die Rolle von psychischer Belastung, Lernunterstützung und veränderten Arbeitsbedingungen der Eltern

Markus Vogelbacher[1,2], **Thorsten Schneider**[1]
[1]Universität Leipzig; [2]Leibniz-Institut für Bildungsverläufe, Bamberg

Zu Beginn der Coronapandemie gehörten Schulschließungen, Kurzarbeit und Homeoffice zu wichtigen Maßnahmen, um die Ausbreitung des Virus in Deutschland einzudämmen. Eltern mussten sich parallel zur Dynamik in ihrem Arbeitsumfeld um die Betreuung und Lernunterstützung ihrer Kinder kümmern. Diese mehrfachen Herausforderungen führten oft zu psychischer Belastung der Eltern. Der hier eingereichte Beitrag konzentriert sich auf Kinder in der Grundschule und geht den Fragen nach, ob sich die psychische Belastung der Eltern, die von ihnen bereitgestellte und geleistete Lernunterstützung sowie veränderte Arbeitsbedingung im ersten Lockdown auf die Entwicklung mathematischer Kompetenzen auswirkten.

Studien auf Basis von Trenddaten zeigen für verschiedene Länder Lernverluste durch Schulschließungen, insbesondere bei mathematischen Fähigkeiten und bei jüngeren Schülerinnen und Schülern (Hammerstein et al., 2021; Betthäuser, Bach-Mortensen & Engzel, 2022; Depping et al., 2021). Einige Studien weisen auch eine Zunahme der Leistungslücken entlang der sozialen Herkunft der Schülerinnen und Schüler nach (Engzell et al., 2021; Maldonado & De Witte, 2020). Die psychischen Belastungen und die Lernunterstützung der Eltern variierten während des Lockdowns ebenfalls nach der sozialen Herkunft (Brown et al., 2020; Vogelbacher & Attig, 2022; Bol, 2020). Darüber hinaus waren Homeoffice und Arbeitszeitverkürzungen häufig zu beobachten (Alon et al., 2022). Die bisherige Forschung hat sich nicht systematisch mit den Folgen der psychischen Belastung der Eltern, der Lernunterstützung und der veränderten Arbeitsbedingungen auf die Kompetenzentwicklung der Schüler während des ersten Lockdowns befasst. Diese Themen

bearbeiten wir unter Rückgriff auf das Family Stress Model (Conger et al., 1992) und das Family Investment Model (Conger & Donellan, 2007) und leiten Hypothesen ab.

Die Hypothesen werden mit Daten der Startkohorte 1 des Nationalen Bildungspanels (NEPS) untersucht. Hier wurden Eltern nicht nur vor der Pandemie jährlich befragt und ihre Kinder getestet, sondern auch während der Pandemie. Die Analysen beziehen sich konkret auf 1.512 8-Jährige, fast alle in der zweiten Jahrgangsstufe, die in den Wochen nach der vollständigen Schulschließung im Frühsommer 2020 an Mathematiktests teilgenommen haben. Aus früheren Wellen liegen Testdaten zu mathematischen Fähigkeiten, kognitiven Grundfähigkeiten und Wortschatzkenntnissen vor. Ein Elternteil, in über 95% die Mutter, lieferte Informationen zum eigenen psychischen Wohlbefinden, der Unterstützung des Lernens der Kinder durch die Eltern und zu den veränderten Arbeitsbedingungen während des ersten Lockdowns. Zum psychischen Wohlbefinden und zu lernbezogenen Aktivitäten mit dem Kind liegen zusätzlich Informationen aus Erhebungen vor der Coronapandemie vor. Darüber hinaus liegen umfangreiche soziodemografische Informationen zum Kind und dem Elternhaus vor. Wir analysieren die Forschungsfragen in einem Value-added-Framework mit OLS-Regressionen und auf Basis multipel imputierter Daten.

Im Sinne des Family Stress Models liefern die Regressionsergebnisse Hinweise darauf, dass sich psychische Belastungen während des ersten Lockdowns nachteilig auf die mathematischen Fähigkeiten der Schülerinnen und Schüler auswirkten. Dieser Effekt bleibt auch bei Kontrolle der Pandemie vorangegangener psychischer Belastungen bestehen. Arbeiteten die Eltern pandemiebedingt von zu Hause aus, fallen die Testergebnisse der Schüler ebenfalls schlechter aus. Entgegen der aus der Theorie hergeleiteten Annahme lässt sich nicht nachweisen, dass dieser Effekt durch wahrgenommene psychische Belastungen vermittelt wird. Entgegen dem Family Investment Model zeitigt die elterliche Lernunterstützung während der Schulschließungen keine Effekte auf die gemessenen mathematischen Fähigkeiten (jedoch Investitionen in Lernförderung vor der Corona-Pandemie). In den Fällen, in denen das befragte Elternteil die Arbeitszeit reduzierte, fallen auch die Testergebnisse des Kindes höher aus. Dies steht in Einklang mit beiden theoretischen Ansätzen.

Eltern-Lehrer:innen-Kooperationen: Unsicherheiten vs. Gelingensbedingungen

Manuela Heimbeck
SRH Heidelberg, Deutschland

Partizipation und Kooperation zwischen pädagogischem Fachpersonal und Eltern können sowohl in der Praxis als auch in der aktuellen Erziehungs- und Bildungsdiskussion als ein von Unsicherheiten, Diskrepanzen und Herausforderungen geprägtes Feld bezeichnet werden (Betz, 2017; Heimbeck, 2020). Die Pandemiesituation fungierte in diesem Zusammenhang als eine Art Katalysator und machte bestehende Missstände in der Eltern-Lehrer:innen-Kooperation (ELK) wortwörtlich “sichtbar(er)“, vor allem beim familiären Umfeld sowie der Lehrkräfteprofessionalisierung.

Es ist hinreichend bekannt, dass das familiäre Umfeld als starker Prädiktor schulischer Lernleistung angesehen werden kann (BMBF, 2021; BMFSFJ, 2020). Die prägenden ersten Grundschuljahre verbrachten einige Schüler:innen größtenteils im Distanzunterricht und waren auf Unterstützungsleistungen von Familie und Schule angewiesen. Jüngste Befunde zeigen, dass viele

Grundschüler:innen währenddessen grundlegende Kompetenzen für eine erfolgreiche Schullaufbahn nicht erworben haben (Voss & Wittwer, 2020; Walper, 2021). Die „Bildungsschere“ zwischen Kindern aus bildungsaffinen Elternhäusern und anderen Milieus wurde nachweislich größer (OECD, 2021). Das führte zu einer verstärkten Forderung eines kollaborativen Miteinanders schulischer und familiärer Akteure. Um diesen Ansprüchen gerecht zu werden, benötigen Lehrkräfte jedoch Kompetenzen, die über das „reine“ Unterrichten hinausgehen und Fähig- und Fertigkeiten wie erziehen, diagnostizieren, beraten, kooperieren und innovieren (Terhart, 2011, 2013). Es ist kritisch anzumerken, dass in der Lehramtsausbildung der Schwerpunkt (noch immer) primär auf dem Unterrichten liegt, was sich u. a. auch in der dritten Lehrerbildungsphase (z. B. Schwerpunkte in der periodischen Beurteilung) widerspiegelt. Diese Fokussierung hat Auswirkungen auf wichtige Gesprächsführungskompetenzen (Scheef & Gleser, 2012) sowie das subjektive Belastungserleben und Unsicherheiten hinsichtlich Kooperation und Kontakt mit Eltern (Rothland, 2013). Der Beitrag untersucht, was eine gelingende Eltern-Lehrer:innen-Kooperation ausmacht und welche Determinanten und Korrelate zum Gelingen oder Scheitern beitragen. Im heuristischen Modell (ELK) werden Befunde zu Informationsfluss, Verhältnis und Lernprozessbegleitung als zentrale Indikatoren identifiziert.

Informationsfluss umfasst alle Kommunikationselemente und -prozesse, die zwischen Schule und Elternhaus stattfinden. Informative Angebote zwischen Eltern und Lehrpersonen stehen im Fokus und erfüllen Kriterien wie Transparenz und Informiertheit (Killus & Paseka, 2021). Das Verhältnis von Schule und Familie beinhaltet insbesondere Prozesse der Beziehungsarbeit zwischen den Akteuren, welche sich mithilfe von Vertrauen und gegenseitigen Erwartungen beschreiben lassen. Hierzu zählen die Adressierung der Akteure, Erwartungen hinsichtlich der Mitwirkungsmöglichkeiten, Chancen der Zusammenarbeit, die Notwendigkeit von Vertrauen sowie individuelle “Bindungs-Typen” im schulischen Kontext (Schneider & Huber, 2020; Schönbächler et al., 2020). Die gemeinsame Lernprozessbegleitung der Schüler:innen durch Eltern und Lehrpersonen erfordert Informationen zum Leistungsstand der Lernenden, zielt aber vor allem auf anwendbare Handlungsimpulse und passgenaue Unterstützungsangebote ab (Porsch & Porsch, 2020; Voss & Wittwer, 2020). Es besteht die Herausforderung, Lernprozesse so zu gestalten, dass subjektive, teils sehr heterogene Lebenswelten Berücksichtigung finden und schulische und außerschulische Lernprozesse sinnvoll miteinander verbunden werden.

Für eine evidenzbasierte Erforschung dieser Faktoren ist ein Mixed Methods Design geplant. Im ersten Teil des Forschungsvorhabens werden qualitative Interviews mit je ca. 15 Erziehungsberechtigten und Lehrer:innen durchgeführt. Die strukturierte qualitative Inhaltsanalyse soll die drei theoretisch postulierten Faktoren induktiv und deduktiv schärfen. Im zweiten Teil des methodischen Vorgehens werden aus den Unterkategorien Fragebogenitems entwickelt. Dieser ELK-Fragebogen soll mit rund 300 Lehrer:innen und Erziehungsberechtigten in einer Querschnittstudie evaluiert werden. Die daran anschließende konfirmatorischen Faktorenanalyse soll prüfen, ob sich die aus der qualitativen Auswertung ergebenen latenten Hauptkategorien auch quantitativ durch entsprechende Faktorenladungen nachweisen lassen. Wie schon etablierte Angebots-Nutzungs-Modelle (u. a. Helmke, 2010) für die Erklärung schulischer Lernleistung komplexe Wirkungsgeflechte

konstatieren, ist auch für die ELK von vielfältigen Wechselwirkungen auszugehen. Für die Indikatoren werden Korrelate auf Ebene der Rechtslage (Makro), auf Einzelschulebene (Meso) und auf Individualebene (Mikro) konstatiert.

Auch wenn diese Faktoren für Praktiker:innen nicht neu sein dürften, erlaubt dieses Modell Ansatzpunkte für das Weiter- und Mitdenken von Maßnahmen zur Verbesserung und Unterstützung einer sinnstiftenden Eltern-Lehrer:innen-Kooperation, mit Ausblick, wie das ELK-Modell mediengestützt umgesetzt werden könnte.

B-02: KI und online Lernen

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S06 S00 B32
Chair der Sitzung: Philipp Marten

Einstellung und Bedenken von Eltern und Lehrkräften gegenüber dem Einsatz von KI im schulischen Kontext

Janne Paula Mesenhöller, Katrin Böhme
Universität Potsdam, Deutschland

Neue Bildungstechnologien wie KI-basierte adaptive Lern-Apps bieten die Möglichkeit, schulische Lerngelegenheiten für heterogene Lerngruppen gewinnbringend zu gestalten. Sie ermöglichen es, individualisierte Lernangebote zur Verfügung zu stellen (Holmes et al., 2018; Walter & Dexel, 2020) und können damit die Lehrkraft bei ihren pädagogischen Tätigkeiten unterstützen. Dies ist für alle Schüler:innen von Vorteil, stellt aber besonders für den inklusiven Unterricht ein großes Potenzial dar (KMK, 2016; Schaumburg, 2020, 2021a, 2021b; Schmid et al., 2021).

Nach Nistor ist *Akzeptanz* „die erste Voraussetzung für die Nutzung einer Bildungstechnologie“ (Nistor, 2020) und damit eine zentrale Rahmenbedingung für den Einsatz von KI-basierten Systemen im schulischen Kontext. Ein gängiges Modell zur Erhebung der Akzeptanz einer Technologie ist das *Technology Acceptance Modell* (TAM, Davis, 1989; Šumak et al., 2011), welches sich für Lehrkräfte und Eltern im Primarbereich als valide erwiesen hat (Scherer & Teo, 2019; Tsuei & Hsu, 2019). Dem TAM zufolge wird die Einstellung gegenüber der Nutzung einer Technologie primär von der *wahrgenommenen Nützlichkeit* (WN) sowie der *wahrgenommenen Benutzerfreundlichkeit* (WB) der Anwendung beeinflusst (Davis, 1989). Basierend auf der *Theory of Planned Behavior* (Ajzen, 1991) und der *Theory of Reasoned Action* (Fishbein & Ajzen, 1975) ist dem TAM zufolge ferner die Verhaltensintention (VI), eine Technologie zu nutzen, der stärkste Prädiktor für die tatsächliche Nutzung.

Trotz der guten empirischen Fundierung des TAM fehlen bislang Befunde zur sozialen Akzeptanz von KI-basierten adaptiven Lern-Apps in der Schule durch die relevante Akteure (bspw. Eltern, Lehrkräfte). Um dieses Forschungsdesiderat zu bearbeiten, wurde im Rahmen des BMBF-geförderten Projekt AKILAS (Adaptiver KI-Lern-Assistent im schulischen Kontext) auf Basis des TAM der Fragebogen *Einstellung von Lehrkräften und Eltern gegenüber KI in der Schule* (ELEKIS) entwickelt. Die Soziale Akzeptanz wurde durch 14 bzw. 12 geschlossene Items (für Lehrkräfte bzw. Eltern) durch eine vierstufige Likertskala operationalisiert. Die postulierte dreifaktorielle Struktur des Fragebogens (WN, WB, VI) wurde mithilfe einer Faktorenanalyse mit dem WLSMV-Schätzer, welcher aufgrund der Verletzung der Normalverteilungsannahme gewählt wurde (Li, 2015), untersucht. Für beide Gruppen erwies sich das dreifaktorielle Modell anhand der gängigen Fitindizes gegenüber einem ein- und zweifaktoriellen Modell als überlegen.

Sowohl für die 141 befragten Lehrkräfte (67.38% weiblich, 31.91% männlich, 0.71% nicht-binär) als auch für die 223 befragten Eltern (76.23% weiblich, 23.32% männlich, 0.45% nicht-binär) konnte entgegen vorheriger Annahmen gezeigt werden, dass die soziale Akzeptanz von KI in der Schule hoch ausgeprägt ist ($\bar{x}$ Lehrkräfte = 2.96, SD = .53, $\bar{x}$ Eltern = 2.94, SD = .65).

In zwei zusätzlichen offenen Items wurden die Lehrkräfte und Eltern ferner explizit nach ihren *Bedenken und Sorgen* hinsichtlich des Einsatzes von KI in der Schule sowie nach gewünschten

Informationen und Maßnahmen gefragt. Mithilfe eines Kodierleitfadens wurden durch zwei unabhängige Kodiererinnen die 1112 Antworten der Eltern und Lehrkräfte in MAXQDA ausgewertet. Die errechnete Interraterreliabilität liegt im zufriedenstellenden Bereich. Die inhaltliche Auswertung des ersten Items zeigt, dass von Eltern und Lehrkräften am häufigsten die Sorge vor nicht ausreichendem Datenschutz (Lehrkräfte 26.4%, Eltern 17%) sowie die Sorge vor einem Rückgang an sozialer Interaktion genannt wurden (Lehrkräfte 15.0%, Eltern 25.1%). Die Lehrkräfte äußerten ferner Bedenken hinsichtlich der aktuellen technischen Voraussetzungen (17.1%) (z.B. fehlendes oder mangelhaftes WLAN, unzureichende Ausstattung mit Hardware). Im zweiten Item wurden ausreichend Datenschutz, funktionierende technische Ausstattung für alle Beteiligten aber auch Weiter- und Fortbildungen (Lehrkräfte 31.9%, Eltern 30.7%) am häufigsten als gewünschte *Informationen und Maßnahmen* genannt.

Basierend auf diesen Ergebnissen wird gegenwärtig das Informationsmodul *Mehr Informationen zum Einsatz von KI in der Schule* (MeInE KI-Schule) entworfen, in dem Eltern und Lehrkräften grundlegende Informationen zum Thema KI vermittelt, Potenziale dargelegt und realistische Grenzen von intelligenten Systemen skizziert werden sollen. Ziel ist es, durch diese transparente Wissensvermittlung die Potenziale von adaptiven KI-basierten Lernunterstützungstools im schulischen Kontext nutzbar zu machen.

Stärkung der Resilienz von Jugendlichen gegenüber Online-Falschinformationen durch ein strategieorientiertes Training

Philipp Marten, Sandra Aßmann, Marc Stadtler
Ruhr-Universität Bochum, Deutschland

Theoretischer Hintergrund & Ziele

Mit wenig Aufwand können im Internet Inhalte gefunden, bearbeitet und geteilt werden. Weil dabei jedoch keine unabhängige Kontrolle stattfindet, müssen Internetnutzer:innen die Gültigkeit der gefundenen Informationen selbstständig bewerten (Metzger et al., 2003).

Zentrale Strategien der Informationsvalidierung sind die Bewertung von Quellen („Sourcing“) sowie das Abgleichen von Aussagen hinsichtlich des Stützungsgrads durch andere Informationsanbieter („Corroboration“). Beide Strategien beherrschen junge Menschen oft nur unzureichend (Wineburg & McGrew, 2019).

Um Informationsvalidierungskompetenzen zu fördern, sind verschiedene Interventionen entwickelt und empirisch erprobt worden (z.B. Bråten et al., 2019; McGrew, 2020; Pérez et al., 2018). Allerdings richten sich diese häufig an ältere Lernende, fokussieren ausschließlich auf die Quellenbewertungsstrategie oder weisen zum Teil methodische Mängel wie das Fehlen von Vergleichstrainings auf.

In der vorliegenden Studie wurden daher die Effekte eines Trainings zur Förderung von Strategien der Informationsvalidierung (Sourcing und Corroboration; *Strategietraining*) mit den Effekten eines Trainings verglichen, das vor allem auf den Erwerb von Sachwissen zum Thema Fake News abzielt (*Wissenstraining*).

Methode

Stichprobe & Design

An der Studie nahmen 196 Jugendliche (M = 13.39 Jahre; 53% weiblich) der 7. und 8. Jahrgangsstufe in einem gemischten Within (3 Messzeitpunkte) * Between (2 Trainingsbedingungen)-Subjects Design teil. Diese wurden im Alfried Krupp-Schülerlabor der RUB klassenweise randomisiert einem *Strategie*- oder *Wissenstraining* zum Thema Online-Falschinformation zugeordnet. Faktoren wie Schulform oder Schul-Sozialindexstufe wurden ausbalanciert.

Beschreibung der Trainings

Strategietraining: Im Mittelpunkt des dreistündigen *Strategietrainings* steht die Förderung der für die Informationsvalidierung zentralen Strategien Sourcing und Corroboration. Diese werden anhand von authentischen Materialien diskutiert und in Gruppen vertiefend eingeübt.

Wissenstraining: Das ebenfalls dreistündige *Wissenstraining* vermittelt ein theoretisches Verständnis des Phänomens Online-Falschinformation mit einem Fokus auf kognitionspsychologische Grundlagen der Verarbeitung von Falschinformationen (Britt et al., 2019). Informationsvalidierungsstrategien werden nicht geübt.

Versuchsablauf & Messinstrumente

Beide Trainings wurden als Workshoptag durch den Erstautor durchgeführt. Zu drei Messzeitpunkten (unmittelbar vor dem Training, unmittelbar danach und drei bis vier Wochen nach Trainingsteilnahme) wurden die folgenden Konstrukte webbasiert erhoben:

(a) Fähigkeit, zwischen Quellen höherer und geringerer Vertrauenswürdigkeit zu unterscheiden (30 Items, Cronbachs Alpha = .682)
(b) Fähigkeit, Falschaussagen durch Anwendung der Corroboration-Strategie zu entlarven (12 Items, Cronbachs Alpha = .772)
(c) Wissen über Strategien der Informationsvalidierung (freie Reproduktion)

Als Kontrollvariablen wurden Lesekompetenz (Schneider et al., 2007), selbsteingeschätztes Vorwissen, Interesse am Thema sowie sozio-ökonomischer Hintergrund der Lernenden erhoben.

Ergebnisse

In Bezug auf die Fähigkeit, zwischen Quellen höherer und geringerer Vertrauenswürdigkeit zu unterscheiden, zeigt eine gemischte ANCOVA mit den Faktoren Messzeitpunkt und Trainingsvariante den erwarteten Interaktionseffekt, $F(2,388) = 9.76$, $p < .001$, $\eta_{part2} = .05$. Während die Kontrollgruppe vor der Workshopteilnahme signifikant besser abschneidet, unterscheiden sich die beiden Trainingsgruppen unmittelbar nach dem Workshop nicht mehr. Mehrere Wochen nach dem Training gibt es wiederum statistisch signifikante Vorteile zugunsten der Strategietrainingsgruppe.

Auch in Bezug auf die Fähigkeit, Falschinformationen mittels der Corroboration-Strategie zu entlarven, zeigt sich eine signifikante Interaktion zwischen den Faktoren Messzeitpunkt und Trainingsvariante, $F(2,340) = 12.38$, $p < .001$, $\eta_{part2} = .07$. Vor Workshopteilnahme schneidet die Kontrollgruppe im Corroboration-Test besser ab. Direkt und mehrere Wochen nach dem Workshop weist jedoch die Strategietrainingsgruppe eine signifikant bessere Leistung auf.

Ein ähnliches Ergebnismuster zeigt sich bei der freien Reproduktion von Informationsvalidierungsstrategien mit signifikanter Interaktion zwischen den Faktoren Messzeitpunkt und Trainingsvariante, $F(2,388) = 5.379$, $p = .005$, $\eta part2 = .03$. Vor Workshopteilnahme unterscheiden sich beide Trainingsgruppen nicht, unmittelbar und mehrere Wochen nach dem Workshop schneidet die Strategietrainingsgruppe signifikant besser ab.

Diskussion

In einer sich immer schneller verändernden Welt bieten flexibel einsetzbare Strategien zur Überprüfung von Internetinformationen einen Bewältigungsmechanismus für die zunehmende Komplexität. Diese Kompetenzen müssen jedoch trainiert werden, damit der angemessene Umgang mit Online-Falschinformationen reflektiert stattfindet. In dem vorgestellten *Strategietraining* erwerben Jugendliche einen Methodenkasten, der den Erwerb von Kompetenzen der Informationsvalidierung kurz- und mittelfristig fördert.

Künstliche Intelligenz in institutionellen Bildungssettings – eine systematische Übersicht aus pädagogisch-didaktischer Perspektive

Verena Köstler
Universität Passau, Germany

Neben digitalisierungsbezogenen Fragestellungen im Kontext Lehren und Lernen im Allgemeinen rücken in jüngster Zeit, auch im deutschsprachigen Raum, mit Künstlicher Intelligenz (KI) verbundene Fragen im Besonderen in den Fokus der Forschung, die zum jetzigen Zeitpunkt noch mit großer Unsicherheit behaftet sind (vgl. bspw. Zehner, 2019). Auch wenn KI bereits viele Lebensbereiche durchdringt, wird das Themenfeld für Bildungsfragen bislang wenig systematisch aufgegriffen, sei es als medienpädagogischer Gegenstand oder – wie im Fokus dieses Beitrags – als mediendidaktischer Bestandteil bei der Gestaltung von Lehr-Lern-Gelegenheiten. In der weitgehend interdisziplinären Bearbeitung, international aufgegriffen unter *artificial intelligence in education (AIED)*, sind technologisch geprägte Fachdisziplinen im Vergleich zu pädagogisch-didaktischen Fachdisziplinen weitaus breiter vertreten (vgl. bspw. für die Hochschuldidaktik: Zawacki-Richter et al. 2019). Eine gemeinsame Herausforderung von KI-Forschung zu Bildungsprozessen und erziehungswissenschaftlicher (Unterrichts-)Forschung besteht in der Erprobung im pädagogischen Feld. Während allerdings erziehungswissenschaftliche Forschung genuin auf konzeptuellen Überlegungen beruht, die von der lernenden Person in einem komplexen Interaktionsgeschehen ausgehen, setzt im Bereich *artificial intelligence in education* erst in jüngster Zeit ein Prozess des Umdenkens ein, die unter den Begriffen *hybrid intelligence paradigm* und *hybrid human-AI-frameworks* (Holstein et al., 2020; Molenaar, 2020) firmieren.

Zur Stärkung eines erziehungswissenschaftlichen Zugangs zu KI-bezogenen Fragestellungen soll ein systematischer Überblick zu KI-gestützten Lehr-Lern-Szenarien in institutionellen Bildungssettings erstellt und ergründet werden, inwiefern (ggf. auch latente) lehr-lern-theoretische Einbettungen vorzufinden sind, welche konkreten didaktischen Intentionen mit ihnen verbunden sind und ob bzw. welche kritische Reflexionen im Hinblick auf institutionelle Implementationen angeführt werden. KI-gestützte Lehr-Lern-Szenarien aus pädagogisch-didaktischer Perspektive herauszuarbeiten trägt dazu bei, Evidenz zu generieren, die einen Beitrag zur Klärung der Rolle

erziehungswissenschaftlicher Disziplinen bei der interdisziplinären Auseinandersetzung mit KI-bezogenen Bildungsfragen leistet.

Methodisch wurde auf ein systematisches Reviewverfahren zurückgegriffen, das auch bei der Bearbeitung erziehungswissenschaftlicher Fragestellungen zunehmend an Bedeutung gewinnt (vgl. Zawacki-Richter et al 2020). Der Prozess von der Entwicklung der Forschungsfrage, die in diesem Projekt darauf abzielte, ein sich im Entstehen befindendes Forschungsfeld zu umreißen, bis hin zur Ergebnisdarstellung, umfasste mehrere Schritte (vgl. Gough et al., 2017; Newman & Gough, 2020). Grundsätzlich war es zunächst erforderlich, einen möglichst repräsentativen Studienkorpus durch eine geeignete Suchstrategie sowie die Spezifizierung von In- und Exklusionskriterien zu generieren. Kennzahlen der Sensitivität und Präzision verdeutlichten, wie gut es in einem von iterativen Schleifen geprägten Suchprozess gelang, für die Forschungsfrage relevante Beiträge zu identifizieren (Best et al, 2014, 351). In der sich anschließenden Synthese wurde der generierte Studienkorpus im Hinblick auf Zielkategorien der übergeordnete Forschungsfrage analysiert.

Diesem Beitrag lag dabei eine datenbankbasierte Suchstrategie zu Grunde. Diese umfasste die Datenbanken ERIC (via Proquest), FIS Bildung (via Fachportal Pädagogik), Learntec sowie Web of Science (via Clarivate). Ein sich schlussendlich auf zwei Wortgruppen („Künstliche Intelligenz“, „Adressaten institutioneller Lehr-Lernsettings“) fokussierender Suchstring wurde entwickelt, mit dem für einen Zeitraum ab 2020 insgesamt 442 zu kodierende Treffer vorlagen, von denen nach Anwendung von In- und Exklusionskriterien in zwei Screeningprozessen (Abstract und Titel, Volltext) ein finaler Studienkorpus mit 51 Beiträgen vorlag. Die Sensitivität (der Datenbanken) reichte dabei von zwei bis 80 Prozent, die Präzision umfasste eine Bandbreite von zwei bis 17 Prozent.

Zum aktuellen Zeitpunkt befindet sich die Synthese des Studienkorpus kurz vor dem Abschluss. Auf der Tagung sollen der Prozess der Durchführung des systematischen Reviews, insbesondere der Generierung des Studienkorpus sowie die Ergebnisse in Form eines systematischen Überblicks zu KI-gestützten Szenarien in institutionellen Bildungssettings vorgestellt und im Hinblick auf die Rolle erziehungswissenschaftlicher Disziplinen im interdisziplinären KI-Diskurs diskutiert werden. Dies schließt Herausforderungen mit ein, die sich durch eine von technologischen Perspektiven dominierte Interdisziplinarität ergeben als auch solche, die operative Aspekte im Prozess der Durchführung aufgreifen.

Digitale Lernangebote zur Steigerung der Medien- und Gesundheitskompetenz: Die Entwicklung und Wirksamkeit einer evidenzinformierten Onlineintervention und die zentrale Bedeutung der User Experience

Lars König, Tim Hamer
Stiftung Gesundheitswissen, Deutschland

Theoretischer Hintergrund / Fragestellung

Gesundheitskompetenz beschreibt die Fähigkeit einer Person, sich mit gesundheitsrelevanten Fragestellungen auseinanderzusetzen und Entscheidungen zu treffen, die dem Erhalt und der Förderung der eigenen Gesundheit dienen (Sørensen et al., 2012). Bei der Suche nach Gesundheitsinformationen wird heutzutage vermehrt auf digitale Medien und das Internet zurückgegriffen, wobei die Qualität und Vertrauenswürdigkeit der verfügbaren Informationen enorm

variiert (Choi & Stvilia, 2015). Um die Gesundheitskompetenz zu fördern, muss dementsprechend auch die Medienkompetenz von Informationssuchenden geschult werden, damit diese zwischen vertrauenswürdigen und weniger vertrauenswürdigen Informationen unterscheiden können. Aus repräsentativen Erhebungen ist bekannt, dass die Mehrheit der Deutschen über eine nur eingeschränkte Gesundheitskompetenz und eine eingeschränkte Medienkompetenz verfügt (Schaeffer et al., 2021). Folgerichtig werden vermehrt Interventionen zur Förderung der Gesundheits- und Medienkompetenz entwickelt. Zunehmend werden in diesem Kontext auch Onlineinterventionen genutzt. Oftmals wird bei der Entwicklung dieser Onlineinterventionen allerdings die User Experience außer Acht gelassen. Ziel der vorliegenden Arbeit war es zu untersuchen, inwieweit eine frei verfügbare und evidenzinformierte Onlineintervention die Gesundheits- und Medienkompetenz fördern kann. Zudem wurde untersucht, inwieweit verschiedene Dimensionen der User Experience den Erfolg der Onlineintervention beeinflussen, um diese Dimensionen bei zukünftigen Angebotsentwicklungen adäquat zu berücksichtigen.

Methode

323 Schülerinnen und Schüler aus allen Bundesländern (*M*=17.88 Jahre, *SD*=1.43, 188 weiblich, 132 männlich, 3 divers) machten demografische Angaben und beantworteten eine Fragebogenbatterie (Prä-Messung). Die Fragebogenbatterie beinhaltete u.a. Instrumente zur Erfassung der Gesundheits- und Medienkompetenz (z.B. Health Literacy for School-Aged Children Instrument) (Paakkari et al., 2016), der wahrgenommen Lehrqualität (z.B. Student Evaluation of Educational Quality Questionnaire) (Daumiller et al., 2021) und des Wissensstandes. Innerhalb von sieben Tagen absolvierten die Schülerinnen und Schüler die bereitgestellte Onlineintervention. Anschließend wurde u.a. die User Experience anhand der Dimensionen (1) Einfachheit, (2) Vielseitigkeit, (3) Farbigkeit und (4) Kunstfertigkeit des Visual Aesthetics of Websites Inventory Instruments (Moshagen & Thielsch, 2010) erhoben und die Schülerinnen und Schüler beantworteten die Fragebogenbatterie erneut (Post-Messung).

Ergebnisse

Es wurde die Hypothese aufgestellt, dass die Schülerinnen und Schüler nach der Absolvierung der Onlineintervention eine höhere Gesundheits- und Medienkompetenz und einen höheren Wissensstand aufweisen würden. Für die gerichtete Hypothesenprüfung wurden Bayessche t-Tests für abhängige Stichproben berechnet (van Doorn et al., 2021). Die Ergebnisse liefern starke Evidenz für die Hypothese. Die Studienteilnehmer erreichten beispielsweise höhere Kompetenzniveaus in den Bereichen Theoretisches Wissen (BF-0=676000; δ=-0.316), Praktisches Wissen (BF-0=92300; δ=-0.294), Kritisches Denken (BF-0=7.42e+13; δ=-0.482), Selbsterkenntnis (BF-0=11500000; δ=-0.345) und Bürgersinn (BF-0=266000; δ=-0.306) des Health Literacy for School-Aged Children Instruments. Zudem wiesen die Studienteilnehmer einen höheren Wissensstand (BF-0=3.82e+97; δ=-1.758) im Bereich der (digitalen) Gesundheitskompetenz auf. Zudem wurden zwei Bayessche Multiple Regressionsanalysen mit den Kriteriumsvariablen Lernerfolg und wahrgenommene Lehrqualität berechnet (van Doorn et al., 2021). Als Prädikatorvariablen dienten jeweils die vier Dimensionen des Visual Aesthetics of Websites Inventory Instruments. Die Ergebnisse sprechen zur Vorhersage des

Lernerfolgs für ein Regressionsmodell, welches die Einfachheit und Vielseitigkeit berücksichtigt (BF10=1.246e+6; R2=0.110). Zur Vorhersage der wahrgenommenen Lehrqualität sprechen die Ergebnisse für ein Regressionsmodell, welches die Vielseitigkeit berücksichtigt (BF10=17900.868; R2=0.074).

Schlussfolgerungen / Bedeutung

Die Ergebnisse sprechen für die Wirksamkeit der frei verfügbaren Onlineintervention zur Förderung der Gesundheits- und Medienkompetenz. Da aufgrund des gewählten Studiendesigns keine Kausalschlüsse möglich sind, sollten in einem nächsten Schritt eine randomisierte kontrollierte Studie zur Überprüfung der Wirksamkeit durchgeführt werden. Lehrerinnen und Lehrern steht mit der Onlineintervention ein niedrigschwelliges, frei verfügbares und wissenschaftlich evaluiertes Angebot zur Verfügung, welches sie zur Förderung der Gesundheits- und Medienkompetenz einsetzen können. Das Angebot der Onlineintervention wird kontinuierlich aktualisiert, erweitert und wissenschaftlich evaluiert. Zudem sollte bei der Entwicklung von Onlineinterventionen neben dem Inhalt auch die User Experience berücksichtigt werden. Insbesondere sollten die Einfachheit und die Vielseitigkeit von Onlineinterventionen adressiert werden, da diese sich zur Vorhersage des Lernerfolgs und der wahrgenommen Lehrqualität eignen.

B-03: COVID-19 Pandemie I

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S06 S00 B41
Chair der Sitzung: Magdalena Stacheder

Auswirkungen der COVID-19-Pandemie auf die sozial-emotionalen Kompetenzen und das Wohlbefinden von Kindern - ein Ländervergleich.

Magdalena Stacheder, Elisa Oppermann, Yvonne Anders
Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund

Der Erwerb von sozial-emotionalen Kompetenzen gehört zu den wichtigsten Entwicklungsaufgaben des Kindesalters und umfasst Fähigkeiten wie Emotionsregulation oder Perspektivübernahme (Petermann & Wiedebusch, 2016; Scheithauer et al., 2016) . Empirische Befunde weisen darauf hin, dass sozial-emotionale Kompetenzen für den künftigen zwischenmenschlichen und akademischen Erfolg entscheidend sind (Moffit et al., 2011). Ferner zeigen sich Zusammenhänge mit Wohlbefinden sowie psychischer Gesundheit (Denham et al., 2009; Garner, 2010). Die Familie ist für die Entwicklung dieser Kompetenzen der erste Kontext (Roßbach et al., 2008), wobei eine hohe Qualität in der häuslichen Lernumwelt mit weniger Problemverhalten (Foster et al., 2005) und mehr sozialen Kompetenzen einhergeht (Rose et al., 2018). Neben der Familie stellt die Kindertageseinrichtung einen wichtigen Entwicklungskontext dar, wobei insbesondere die Qualität früher Bildung und Erziehung entscheidend für die kindliche sozial-emotionale Entwicklung ist (z.B. Broekhuizen et al., 2016). Auch die Interaktion mit Gleichaltrigen ist essentiell, um sozial-emotionale Kompetenzen zu erlernen (Parke & Clarke-Steward, 2011). Die Entwicklung der sozial-emotionalen Kompetenzen ist folglich in soziale Interkationen in verschiedenen Lernumgebungen eingebettet (Bronfenberenner & Morris, 2006), die durch die Covid-19 Pandemie massiv beeinflusst wurden. Viele Länder beschlossen restriktive Maßnahmen, um die Ausbreitung des Virus zu verlangsamen. Während einige Länder, wie beispielsweise Deutschland, strenge Restriktionen (z.B. Schließungen von Bildungsinstitutionen und Ausgangssperrungen) einführten, wurde in anderen Ländern, wie beispielsweise Norwegen, oftmals lediglich Empfehlungen ausgesprochen (Ritschie et al., 2022). Da solche Restriktionen die natürlichen Entwicklungskontexte für sozial-emotionale Fähigkeiten der Kinder in unterschiedlicher Weise einschränken, lässt sich eine Auswirkung der Covid-19-Pandemie auf die sozial-emotionalen Kompetenzen und das Wohlbefinden vermuten, die sich in Abhängigkeit der Strenge der Maßnahmen unterscheiden.

Fragestellung

Das Ziel des vorliegenden Reviews ist es, Auswirkungen der Covid-19 Pandemie auf die sozial-emotionalen Kompetenzen und das Wohlbefinden von Kindern im Alter von null bis zwölf Jahren zu identifizieren sowie die Ergebnisse mit der Strenge der restriktiven Maßnahmen in den unterschiedlichen Ländern in Verbindung zu bringen.

Methode

Zur Beantwortung der Fragestellung wurde eine systematische Suche in acht verschiedenen Datenbanken (u.a. ERIC und PsycInfo) durchgeführt, wofür deutsche und englische Suchbegriffe zu sozial-emotionalen Kompetenzen und zum Wohlbefinden (z.B. wellbeing) mit Suchbegriffen bezüglich der Covid-19 Pandemie (z.B. Covid-19) und mit kindbezogenen Suchbegriffen (z.B. child) kombiniert

wurden. Die Suche ergab 8923 Treffer, nach Entfernung der Duplikate blieben 7441 Studien. Die vorläufige Anzahl der Studien nach Screening von Titel und Abstract umfasst 898. Um für die Forschungsfrage relevante Studien zu identifizieren, wurden Ein- und Ausschlusskriterien definiert (z.B. Ausschluss von physischen Auswirkungen wie Schlafqualität). Anhand der Maßnahmen in den jeweiligen Ländern, wurden diese zu den Kategorien „sehr streng“, „streng“ und „mild“ zugeordnet.

Erste Ergebnisse und ihre Bedeutung

Vorläufige Auswertungen der ersten Studien (N = 28), basierend auf Daten mit Messzeitpunkten vor sowie während der Pandemie, weisen eine gemischte Befundlage auf. So zeigen die Ergebnisse einiger Studien Verschlechterungen (vorläufiges N=10, u.a. Feinberg et al., 2022), andere keine Veränderungen (vorläufiges N=4, u.a. Linnavalli & Kalland, 2021) in den kindlichen sozial-emotionalen Kompetenzen.. Auch bezüglich des kindlichen Wohlbefindens zeigt sich ein inkonsistentes Bild: Es werden sowohl Verbesserungen (vorläufiges N=1, Gray, 2020), keinen Veränderungen (vorläufiges N=4, u.a. Janssen et al., 2020) sowie von Verschlechterungen (vorläufiges N=11, u.a. Ravens-Sieberer et al., 2021) berichtet. Auf Grundlage dieser heterogenen Ergebnisse, wird neben der Auswertung weiterer Längsschnittuntersuchungen im nächsten Arbeitsschritt der Einbezug der Strenge der restriktiven Maßnahmen auf Länderebene vorgenommen. Die Auswertungen aller Studien wird bis Endes des Jahres abgeschlossen sein. Das Review trägt zum besseren Verständnis der Auswirkungen der Covid-19 Pandemie sowie der Restriktionen zur Eindämmung der Pandemie auf die sozial-emotionalen Kompetenzen und das Wohlbefinden von Kindern bei. Der Ländervergleich ermöglicht zudem, die Konsequenzen unterschiedlicher Restriktionen für die kindliche Entwicklung abzuschätzen und hat hohe praktische sowie politische Relevanz, insbesondere mit Blick auf zukünftige gesellschaftliche Herausforderungen.

Schule in Pandemiezeiten - Einblicke in die Wahrnehmung von COVID-19 bedingten Hygienemaßnahmen in Österreichs Schulen durch Lehrer*innen

Flora Woltran[1], Katharina-Theresa Lindner[1], Susanne Schwab[1,2]
[1]Zentrum für Lehrer*innenbildung, Universität Wien, Österreich; [2]Institut für Bildungswissenschaft, Universität Wien, Österreich

Aufgrund der zu Beginn des Jahres 2020 aufgetretenen COVID-19-Pandemie stürzten Bildungssysteme auf der ganzen Welt in eine noch nie dagewesene Krise. Als rasche Reaktion auf die mit dem SARS-CoV-2-Virus verbundenen Gesundheitsgefahren wurden Schulen in mehr als 190 Ländern innerhalb weniger Wochen teilweise oder vollständig geschlossen (UNESCO, 2020). Auch in Österreich wurden Schulschließungen erstmals im Frühjahr 2020 und neuerlich im November 2020 als Mittel zum Zweck der raschen Eindämmung des Infektionsgeschehens und zum Schutze der Schüler*innen und des schulischen Personals eingeführt. Neben der bildungspolitischen Strategie des temporären Aussetzens des Präsenz-Unterrichts wurden regelmäßige COVID-19 Testungen und das Tragen eines Mundnasenschutzes in Klassen oder im gesamten Schulraum implementiert.

Während das Tragen eines Mundnasenschutzes eine vergleichsweise einfach umzusetzende Maßnahme in Schulen ist (Spitzer, 2020) und sich hinsichtlich der Eindämmung von COVID-19-Infektionen als wirksam erwiesen hat (Prather et al., 2020), müssen insbesondere innerhalb des schulischen Kontexts potentielle sozio-emotionale Nachteile mitberücksichtigt werden. So zeigt sich beispielsweise, dass das Tragen eines Mundnasenschutzes verbale und nonverbale Formen der

Kommunikation und emotionale Signale zwischen Lehrer*innen und Schüler*innen stören kann (Spitzer, 2020). Dies wiederum kann negative Folgen für die Beziehung zwischen Lehrer*innen und Schüler*innen, den Gruppenzusammenhalt und die Lernentwicklung mit sich bringen (Spitzer, 2020; Vergara et al., 2022).

Die aufgrund der Pandemie flächendeckend eingeführten COVID-19-Testungen an Schulen gelten als nützliche Strategie zur schnellen Identifizierung und Isolierung infizierter Personen (Krishnaratne et al., 2020). Einige Befunde legen nahe, dass COVID-19 Testungen von Schulleitungen und Lehrkräften als wirksame Maßnahme gegen die unkontrollierte und schnelle Ausbreitung von COVID-19-Infektionen wahrgenommen werden (z.B.: Lorenc et al., 2021; Unger et al., 2021). Allerdings zeigt sich auch, dass COVID-19-Testungen umfangreiche Vorbereitungs- und Planungsaufgaben sowie zusätzliche zeitliche und räumliche Ressourcen erfordern und daher eine logistische Herausforderung für Schuldirektoren*innen und Lehrer*innen darstellen können (Asgary et al., 2021).

In Anbetracht der Schwierigkeiten, die mit der Umsetzung wirksamer COVID-19-Hygienemaßnahmen an Schulen verbunden sein können, befasst sich die vorliegende Studie mit folgenden Forschungsfragen: (1) Wie bewerten Lehrkräfte die Durchführung der COVID-19-Tests und das Tragen eines Mundnasenschutzes in ihren Klassen? (2) Gibt es Unterschiede in der Einstellung der Lehrer*innen zu den COVID-19-Testungen je nach Schultyp (Grundschule, Sekundarschule, Sonderschule)? (2) Beeinflusst die Unterrichtserfahrung die Zufriedenheit der Lehrer*innen mit den COVID-19-Testungen? (3) Lässt sich ein Zusammenhang zwischen den COVID-19-Testungen und der Arbeitszufriedenheit der Lehrer*innen herstellen?

Zur Beantwortung der Forschungsfragen wurden sowohl quantitative als auch qualitative Daten verwendet, die während des dritten Messzeitpunkts der INCL-LEA-Studie (INCLusive Home LEArning) im November 2021 erhoben wurden. Die Datenerhebung umfasste sowohl einen quantitativen Online-Fragebogen (N = 1.170) als auch leitfadengestützte Interviews (N = 5). Der Online-Fragebogen umfasste sowohl geschlossene als auch offene Fragen zu den von den Lehrern wahrgenommenen Chancen und Herausforderungen sowie zu ihrer allgemeinen Einschätzung der aktuellen Arbeitsbedingungen in den Schulen. Die Leitfadeninterviews enthielten Fragen zu den subjektiv wahrgenommenen Gesamterfahrungen der Lehrer*innen während der durch COVID-19 gekennzeichneten Schuljahre und zu ihren Eindrücken von den in Schulen implementierten Maßnahmen.

Vorläufige Ergebnisse zeigen, dass die Durchführung der COVID-19-Testungen im Pflichtschulbereich (Grund- und Mittelschulen) tendenziell gut funktionierte. Im Gegensatz dazu galten die COVID-19 Testungen in den Sonderschulen als problematische Angelegenheit. Darüber hinaus deuten die Befunde darauf hin, dass die COVID-19-Testungen für Lehrer*innen mit langjähriger Berufserfahrung reibungsloser abliefen als für jene mit weniger Berufserfahrung. Auch legen die Ergebnisse nahe, dass sich etwa die Hälfte der Lehrkräfte durch diese Aufgabe stark belastet fühlte. Hinsichtlich des Tragens des Mundnasenschutzes zeigen die qualitativen Befunde ein gemischtes Bild. So berichten zwei der fünf befragten Lehrpersonen von einer positiven Wahrnehmung und führten dies auf den damit verbundenen Schutz vor gesundheitlichen Gefahren zurück. Drei weitere

Lehrer*innen geben hingegen an, das Tragen eines Mundnasenschutzes aufgrund zahlreicher sozio-emotionaler und körperlicher Schwierigkeiten als Störfaktor anzusehen.

Welche Faktoren spielen bei der Bewältigung des Übergangs in die Schule bei Kindern während der Covid-19 Pandemie eine Rolle?

Stefanie Pietz, Iris Wortmann, Yvonne Anders
Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund

Der Übergang von Elementarbereich in die Primarstufe steht jährlich für derzeit etwa 770.000 Schulanfängerinnen in Deutschland (Autorengruppe Bildungsberichterstattung 2022) an. Die Verantwortlichkeiten für die Rahmenbedingungen des Übergangs liegen bei den Bundesländern, Grundsätze für eine gelungene Umsetzung legt u.a. der Beschluss der JFMK & KMK 2009 „Den Übergang von der Tageseinrichtung für Kinder in die Grundschule sinnvoll und wirksam machen“. Wildgruber und Niesel (2016) analysieren in ihrer Analyse Anknüpfungspunkte eines erfolgreichen Übergangs. Weitere Studien werfen die Perspektive auf curriculare Abstimmungen (z.B. Kluczniok etal. 2015) oder Kooperationsmaßnahmen zwischen Kindergarten und Grundschule (z.B. Faust etal. 2013). Forschungen um den Transitionsansatz (Griebel, Niesel 2011) liefern Erkenntnisse zur Beteiligung der verschiedenen Akteure (Kinder, Eltern, Fachkräfte). Kritisch diskutiert wird in diesem Ansatz häufig der Fokus auf die Problembelastung hinsichtlich der Übergangsbewältigung (z.B. Faust, etal. 2012)

Die Coronapandemie stellt seit Frühjahr 2019 klassische Kooperationsaktivitäten (z.B. Schulbesuche von Kindergartenkindern, Elternabende) zwischen Elementar-/Primarbereich, Aktivitäten mit Vorschulkindern sowie die Durchführung der verankerten Einschulungsuntersuchungen in den Bundesländern vor neue Herausforderungen. Ausgehend von der paradoxen Theorie (Caspi, Moffit 1993) ist anzunehmen, dass keine problembehafteten Verhaltensweisen oder Persönlichkeitsveränderungen im Rahmen der Übergangsphase hervorgerufen werden. Inwieweit pandemiebedingte Maßnahmen (z.B. Einrichtungsschließungen, Notbetreuung, Kontaktvermeidung) Einwirkungen auf die Schuleintrittsphase und dessen Bewältigung haben ist noch unklar. Zieht man Erkenntnisse des Transitionsansatzes (Griebel, Niesel 2011) heran, wäre die Bewältigung des Schuleintritts und den zusätzlichen Herausforderungen der Coronapandemie mit weiteren Belastungen verbunden.

Fragestellung

Vor dem Hintergrund der Coronapandemie und den Restriktionen wird in der Studie untersucht, welche Faktoren für den Übergang von Kindergarten in die Grundschule bedeutsam sind. Dafür werden zum einen Unterschiede in der Übergangsbewältigung zwischen Kindern, die während des Lockdowns die Kindertageseinrichtung besuchten und Kindern, die zu Hause betreut wurden, untersucht. Zum anderen werden Zusammenhänge zwischen institutionellen/externen Bedingungen (z.B. Schuleingangsuntersuchung) und der Übergangsbewältigung erforscht. Außerdem wird der potenzielle Beitrag von sozialen Aspekte (z.B. prosoziales Verhalten) für eine gelingende Bewältigung des Übergangs überprüft.

Methode

Im Rahmen einer Familienbefragung wurden von November 2021 bis Januar 2022 Eltern standardisiert und onlinebasiert befragt. Die Gesamtstichprobe umfasst 253 Eltern mit Kindern in der ersten bis dritten Klasse (49% 1.Klasse, 33% 2.Klasse, 18% 3.Klasse). Die Studie umfasst die Einschulungskohorte 2021 mit 104 befragten Eltern. Neben kindspezifischen Merkmalen wurden sie zur Notbetreuung durch die Kindertageseinrichtung und zum Einschulungszeitpunkt (fristgerecht/nicht-fristgerecht) ihres Kindes befragt. Des Weiteren wurden Angaben zur Durchführung von Schuleingangsuntersuchungen erfragt (stattgefunden/ausgefallen). Darüber hinaus wurden sie zu den Auswirkungen des Lockdowns auf ihr Kind (*„Mein Kind hat soziale Probleme als Folge des Lockdowns.“,* α = .80), zum prosozialen Verhalten auf Basis des SDQ-Deu (*„Mein Kind teilt gerne mit anderen Kindern.“,* α = .65) und zur Übergangsbewältigung (*„Mein Kind geht gern in die Schule.“*, α = .85) befragt. Es wurden Gruppenvergleiche und erste Zusammenhangsanalysen durchgeführt. Weitere multivariate Analysen sind geplant.

Ergebnisse

Die Ergebnisse zeigen keine statistisch signifikanten Unterschiede in der Übergangsbewältigung zwischen Kindern in der Notbetreuung während des Lockdowns und Kindern, die zu Hause betreut wurden (T=-.03, *p*=.979). Mit Blick auf kindspezifische Merkmale zeigt sich, dass Mädchen der Übergang besser gelingt als Jungen (*r*=.180, *p*=.076) und die Familiensprache keinen statistisch signifikanten Zusammenhang aufweist. Der Übergang in die Grundschule fällt den Kindern leichter, wenn sie fristgerecht eingeschult werden (*r*= -.411, *p*<.001). Kinder, bei denen aufgrund der Coronapandemie keine Schuleingangsuntersuchung stattfand, bewältigen den Schuleintritt genauso gut wie Kinder, die an einer Schuleingangsuntersuchung teilgenommen haben (T=-.30, p=.77). Die Bedeutung sozialer Faktoren scheint für die Schulbewältigung von größerer Relevanz zu sein als institutionelle/externe Bedingungen. Je höher die negativen Auswirkungen des Lockdowns auf das Kind eingeschätzt werden, desto geringer/schlechter ist die Schulbewältigung (*r*= -.36, p<.001). Wird des Weiteren das prosoziale Verhalten positiv eingeschätzt, gelingt die Schulbewältigung besser (*r*= .29, *p*=.003). Die Ergebnisse werden bezüglich des weiteren Forschungsstandes diskutiert.

Die COVID-19-Pandemie im Schuljahr 2020/21: Wie haben sich die kognitiven Grundfähigkeiten von Schulkindern der Primarstufe entwickelt?

Wolfram Rollett[1], Thomas Leitgeb[2], Katja Scharenberg[1]
[1]Pädagogische Hochschule Freiburg, Germany; [2]Pädagogische Hochschule Burgenland

Die pandemieeinschränkenden Maßnahmen hatten weitreichende Auswirkungen auf die Lebenswelt von Kindern und Jugendlichen. Ihre individuellen Lern- und Erfahrungsmöglichkeiten waren in dieser Zeit stark eingeschränkt. Damit stellt sich die Frage, wie sich dies auf ihre Entwicklung ausgewirkt hat. Für schulische Leistungen sind Entwicklungsnachteile für Schülerinnen und Schüler (s. z.B. Schult et al. 2022) wie auch die Ausweitung herkunftsbedingter Schereneffekte empirisch bereits gut belegt (z.B. Juniper Education 2021). Für die Entwicklung kognitiver Grundfähigkeiten fehlen derartige Analysen bisher. Unser Betrag nimmt daher die kognitive Entwicklung von Schulkindern der Primarstufe im Schuljahr 2020/21 in den Blick. Geprüft wird dabei, ob sich die untersuchten Kinder in ihren kognitiven Grundfähigkeiten in dem neun Monate umfassenden Untersuchungszeitraum in dem

Maß positiv weiterentwickeln konnten, wie dies z.B. nach den metaanalytischen Befunden von Rindermann (2011) unter regulären Lebensbedingungen erwartbar gewesen wäre (Fragestellung 1). Darüber hinaus wird Effekten potenziell bildungsbenachteiligender bzw. -bevorteilender individueller Merkmale nachgegangen, die sich im Sinne von Schereneffekten auf die kognitive Entwicklung der Kinder auswirken können (Fragestellung 2).

Methoden und Stichprobe

Datengrundlage ist eine im Schuljahr 2020/21 in der dritten Schulstufe an drei österreichischen Volksschulen (Primarstufe) in fünf Klassen durchgeführte Längsschnittstudie. Die Bruttostichprobe umfasste 119 Schülerinnen und Schüler. Davon nahmen 104 Kinder (87,5 Prozent; 48 Mädchen, 56 Jungen) teil. MZP1 lag zu Schuljahresbeginn im September 2020. Weitere Datenerhebungen erfolgten im Mai 2021 (MZP2) und zum Schuljahresende im Juni 2021 (MZP3). Von 38 Unterrichtswochen erfolgten 21 in Präsenz und 17 mit Fernunterricht. Darüber hinaus kam es zu einschneidenden gesamtgesellschaftlichen Maßnahmen wie „Lockdowns“ mit Ausgangs-, Veranstaltungs- und Versammlungsbeschränkungen (85 Tage).

Abhängige Variablen

Zur Erfassung der kognitiven Grundfähigkeiten der Schülerinnen und Schüler wurden zu MZP1 und MZP3 die Subtests *Schlussfolgerndes Denken* (Retestreliabilität = .67) und *Rechnerisches Denken* (Retestreliabilität = .77) des Kognitiven Fähigkeitstests für Kinder (KFT 1-3; Heller & Geisler 1983) eingesetzt.

Unabhängige Variablen

Die unabhängigen Variablen wurden aus testökonomischen Gründen zu MZP2 erhoben: Geschlecht, Alter, Migrationshintergrund, familiärer schulischer Bildungshintergrund, Anzahl der Bücher zu Hause, Familienwohlstand sowie Vorhandensein von digitalen Endgeräten und Internetzugang.

Fehlende Werte

Fehlende Werte wurden mit *R* (R Core Team 2017) in *RStudio* (RStudio Team 2020) multipel imputiert (Ausfallprozess MAR; MICE, Buuren & Oudshoorn 2011; CART, Graham et al. 2007; 20 Imputationen mit 40 Iterationen).

Auswertungsmethoden

Die Auswertung erfolgte mittels *t*-Test für abhängige Stichproben und Regressionsanalysen (zweiseitige Signifikanztests, $\alpha \leq 5\%$). Die Daten der abhängigen Variablen wurden vor der Durchführung der Regressionsanalysen (Mplus 8.7, MLR, *type = complex,* Muthén & Muthén 1998-2022) am Mittelwert der jeweiligen Klasse normiert.

Ergebnisse

Für keinen der beiden KFT-Subtests ist der für den Untersuchungszeitraum erwartete positive Entwicklungstrend signifikanzstatistisch nachweisbar (Fragestellung 1). Auch hinsichtlich der Streuungen unterscheiden sich die Testleistungen nicht. Für die multiplen Regressionsanalysen wurden alle oben genannten unabhängigen Variablen und die Ausgangsleistung als Prädiktoren herangezogen. Für die Entwicklung im *Schlussfolgernden Denken* und im *Rechnerischen Denken*

erweisen sich neben der jeweiligen Ausgangsleistung zwei Prädiktoren als signifikant: Kinder, deren Eltern einen höheren schulischen Bildungshintergrund haben (β = .175*, p = .011 bzw. β = .163**, p = .004) und in deren Familien eine bessere Ausstattung mit digitalen Endgeräten bzw. einem Internetzugang gegeben ist (β = .332**, p = .004 bzw. β = .106**, p = <.001), entwickeln sich günstiger.

Diskussion

Die Befundlage unterstützt die Annahme, dass für die Gesamtstichprobe in dem von pandemieeinschränkenden Maßnahmen geprägten Untersuchungszeitraum Entwicklungsbedingungen vorlagen, die einer altersgemäßen positiven Weiterentwicklung der kognitiven Grundfähigkeiten entgegenstanden. Zudem deuten die berichteten Befunde darauf hin, dass es in Abhängigkeit von häuslichen Merkmalen zu Schereneffekten gekommen ist. Ob sich die berichteten Befunde verallgemeinern lassen, muss die weitere Forschung zeigen. Sollte dies der Fall sein, wären gezielte Programme angezeigt, um Entwicklungsanreize zu setzen und die erlittenen Nachteile so weit wie möglich auszugleichen.

B-04: Diagnostik und Mathematik

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S06 S02 B06
Chair der Sitzung: Julian Urban

Testing Norms: Comparing Continuous and Conventional Norming Methods for the NFC-KIDS

Julian Urban[1], Vsevolod Scherrer[2], Anja Strobel[3], Franzis Preckel[2]
[1]GESIS—Leibniz Institute for the Social Sciences; [2]Universität Trier; [3]TU Chemnitz

Theoretical background

Psychological tests and questionnaires are routinely used in educational research and practice. These standardized instruments measure psychological constructs decisive for assessment and placement decisions and interventions. An appropriate norming of a standardized instrument is a prerequisite for its purposeful application (Moosbrugger & Kelava, 2020, Urbina, 2014). Usually manuals report within-group norms. For different groups including a specific interval of a norm predictor (e.g., age) inverse normal transformation is applied to extract norms (Lenhard et al., 2019). However, such conventional norms are biased within a group and at the transition between groups (Timmermann et al., 2021). Solving this issue by narrowing the interval leads to an economy-precision dilemma. Continuous norming methods, namely semiparametric and parametric norming (Lenhard et al., 2018; Timmermann et al., 2021), solve these issues economically by applying polynomial regression or the Generalized Additive Models for Locations, Scale, and Shape (Rigby & Stasinopoulos, 2005). This results in smooth functions relating raw and norm scores over a continuous norm predictor.

Continuous norming methods are relatively new but increasingly gain in attention. A systematic literature search using *PsycInfo*, *Web of Science*, and *Eric* revealed 34 articles dealing with continuous norming. Most studies (24) emerged after 2015 and applied continuous norming methods to investigate clinical constructs or intelligence based on real data (25). Some studies compared continuous to conventional norms. Continuous methods excelled if appropriate models were computed (e.g., Lenhard & Lenhard, 2021, Timmermann et al., 2021). However, there is a lack of studies comparing semiparametric with parametric norms.

Research question

This study aimed to close this research gap by comparing conventional, semiparametric, and parametric norms regarding their precision using a real data example of a construct relevant for achievement and motivation, the Need for Cognition (NFC-KIDS; Preckel & Strobel, 2017).

Methods

We examined the original norming sample for the NFC-KIDS of elementary school aged children (N = 2581). Age formed the norm predictor. Besides conventional norms, semiparametric and parametric norms were computed (Gary et al., 2021; Timmermann et al., 2021). For both methods, we compared several models and selected the best fitting models. The absolute fit of these models was assessed visually. The precision analysis of the final models relied on 1000 bootstrap samples and the interpercentile range as indicator for precision (Oosterhuis et al., 2016). This represented the width of a 95% confidence interval of the raw scores associated with a specific norm percentile. Thus, the smaller the width the more precise the norms. Since conventional and semiparametric norms relied on

subgroups, their span was varied to validate the results against effects of sample size per group on precision (e.g., Oosterhuis et al., 2016).

Results and discussion

Model selection resulted in an acceptable absolute fit for both continuous norming methods (i.e., P-P plot, worm plot). The precision analysis revealed a clear hierarchy of the three norming methods. Parametric norms were most precise, followed by semiparametric norms. Conventional norms were least precise. Additionally, an effect of subgroup sample size was also evident. Here, a strong effect occurred on conventional norms, and a small on semiparametric norms indicating a higher precision for fewer groups with a higher sample size. No such effect appeared for parametric norms. The subgroup sample size did not affect the precision hierarchy of norming methods.

To conclude, continuous approaches lead to more precise norms which can be assessed more economically (i.e., smaller sample sizes). To enhance the application of continuous norming methods in educational research and practice, we discuss their possible advantages in predicative validity for educational outcomes as well as the need for a framework suggesting required sample size, norming method, and exact model used.

Der Einfluss von Mindsets auf kognitive Urteilsprozesse von angehenden Lehrkräften beim Analysieren von mehrdeutigen SchülerInnenlösungen zum Bruchvergleich – Eine Experimentalstudie

Elif Özel, Timo Leuders, Katharina Loibl
PH Freiburg, Deutschland

Theoretischer Hintergrund

Korrelationen zwischen Lehrkräfteurteilen und SchülerInnenmerkmalen können zwar die Genauigkeit eines diagnostischen Urteils beschreiben, nicht aber Gründe für Abweichungen aufklären (Südkamp et. al, 2012). Um solche Abweichungen zu erklären, muss der Blick auf die Genese eines Urteils gerichtet werden. Loibl et al. (2020) konzeptualisieren diagnostische Urteile als Informationsverarbeitungsprozesse. Dabei können verschiedene kontextuelle Bedingungen die Wahrnehmung und Verarbeitung von Informationen beeinflussen. Die vorliegende Studie konzentriert sich auf den Einfluss von Mindsets (implemental vs. deliberat). Gollwitzer (2012) beschreibt Mindsets als temporäre „kognitive Orientierungen“ (States), die durch die Situation ausgelöst werden und zur effizienten Lösung dieser Situation die Form der Informationsverarbeitung bestimmen (Achtziger & Gollwitzer, 2010). Einmal aktivierte Mindsets übertragen sich auch auf nachfolgende, nicht verwandte Aufgaben (carry-over effect; Gollwitzer, Heckhausen & Steller, 1990). Dabei zeichnen sich zwei verschiedene Modi der Informationsverarbeitung aus: Das objektive und ungefilterte (open-mindedness) Verarbeiten von Informationen aller Art im deliberaten Modus und das selektive Verarbeiten (closed-mindedness) von Informationen, die die eigene Hypothese bestätigen, im implementalen Modus (Bayer & Gollwitzer, 2005). Neben diesen situativen Modi können auch habituelle Reaktionsweisen (Traits) wie beispielsweise das Bedürfnis nach Deliberationen ähnliche kognitive Prozesse hervorrufen (Betsch, 2004). Weinhuber et al. (2019) konnten erste Erkenntnisse zur Bedeutung von Mindsets auf Lehrkräftehandeln aufzeigen. Bisweilen unbekannt ist der Einfluss von Mindsets und dem Bedürfnis nach Deliberation auf diagnostische Urteilsprozesse.

Forschungsfrage/Hypothesen

Es wird der Frage nachgegangen, ob es **Unterschiede in der Verarbeitung** von diagnostischen Informationen (Anzahl und Art) und der **Urteilsgüte** zwischen angehenden Lehrkräften im deliberaten und im implementalen Modus gibt. Erwartet wird, dass angehende Lehrkräfte im deliberaten Mindset weniger fehleranfällig (Hügelschäfer & Achtziger, 2014) urteilen als angehende Lehrkräfte im implementalen Mindset, da sie eine Vielzahl an Informationen verarbeiten und offen für Informationen aller Art sind.

Bei angehenden Lehrkräften im implementalen Mindset werden schnellere Urteile auf der Basis weniger Informationen erwartet, da sie eine geringere Aufnahmebereitschaft für neue Informationen zeigen. Zudem wird erwartet, dass sie Informationen fokussieren, die ihre Annahmen bestätigen. Aufgrund unzureichender Informationen steigt das Risiko eines Fehlurteils, insbesondere wenn dem SchülerInnenfehler verschiedene Fehlerstrategien zugrunde liegen können.

Methode

Mittels Comics werden angehende Lehrkräfte in den implementalen oder deliberaten Mindset versetzt (vgl. Weinhuber et al., 2019). Anschließend erhalten sie SchülerInnenlösungen zum Bruchvergleich, die sie analysieren sollen. Die Analysen der angehenden Lehrkräfte lassen dabei eindeutige Rückschlüsse darauf ziehen, welche der in der Diagnosesituation enthaltenen Informationen in ihre Urteilsbildung einfließen.

Da sowohl die oben beschriebenen Mindsets (States), als auch das Bedürfnis nach Deliberation (Trait) ähnliche Auswirkungen auf kognitive Informationsverarbeitungsprozesse beim diagnostischen Urteilen haben können, wird Betschs (2004) Persönlichkeitsfragebogen zur Deliberation eingesetzt.

Ergebnisse und ihre Bedeutung

Bisher haben 26 angehende Lehrkräfte an der Studie teilgenommen. Zur Untersuchung von Unterschieden in der Anzahl der verarbeiteten Informationen in den Lehrkräfteurteilen werden lineare gemischte Modelle verwendet mit Mindset und Bedürfnis nach Deliberation als Fixed Effects. Die Modelle beinhalten zudem Random Effects für Lehrkräfte (zur Abbildung individueller Unterschiede zwischen Lehrkräften) und Messzeitpunkte (zur Abbildung individueller Unterschiede zwischen verschiedenen Messzeitpunkten/Vignetten). Die Ergebnisse zeigen, dass das Induzieren von Mindsets einen signifikanten Einfluss ($p = .0005$) auf die Anzahl der Informationen hat, die in den Lehrkräfteurteilen verarbeitet wurden. Hypothesenunterstützend integrierten angehende Lehrkräfte im deliberaten Mindset im Durchschnitt 0,93 Informationen mehr in ihr diagnostisches Urteil (durchschnittlich 2,7 Informationen) als untersuchte Lehrkräfte im implementalen Mindset (durchschnittlich 1,8 Informationen). Das Bedürfnis nach Deliberation (Trait) weist hingegen keinen signifikanten Einfluss auf die Anzahl der genutzten Informationen in den Lehrkräfteurteilen auf ($p = .44$). Der Prädiktor Mindset (State) reduziert im Gesamtmodell die Fehlervarianz von .46 auf .23 und klärt 50 % der zufälligen Varianz zwischen den untersuchten Lehrkräften auf. Die Mindsetinduktion hat somit einen starken Effekt auf die Anzahl der von den Lehrkräften verarbeiteten Informationen.

Die Auswertung zur Urteilsgüte steht noch aus.

Wie gut können Grund- und Förderschullehrkräfte Rechenstörungen identifizieren?

Jil Brambrink, Olga Kunina-Habenicht
TU Dortmund, Deutschland

Theoretischer Hintergrund:

Rechenstörungen sind als spezifische Lernstörung klassifiziert, welche ohne Behandlung bis ins Erwachsenalter bestehen bleiben (APA, 2013). Zudem gehen mit spezifischen Lernstörungen oftmals auch sozial-emotionale Probleme einher, sowie eine erhöhte Neigung zu externalisierenden und internalisierenden Auffälligkeiten (Visser et al., 2019). Besonders wichtig scheint daher eine möglichst frühzeitige Diagnose von spezifischen Lernstörungen, um komorbide psychische Störungen zu vermeiden (Hässler, 2007; Visser et al., 2019) und Förderangebote bereitzustellen.

Neben neurologischen Ursachen scheinen auch inadäquate didaktische Ansätze bei der Entstehung von Lernstörungen eine Rolle zu spielen (Hasselhorn, 2022). Als Teil des Professionswissens ist die diagnostische Kompetenz von Lehrkräften (Baumert & Kunter, 2011) ein wichtiger Indikator für gute Unterrichtsqualität (Voss et al., 2015). Studien zeigten bereits, dass eine höhere diagnostische Sensitivität mit besseren Mathematikleistungen einhergeht (Brunner et al., 2011). Allerdings scheinen die Urteile bei der Einschätzung von Leistungsmerkmalen häufig ungenau zu sein (Spinath, 2005). Doch auch relativ gute Einschätzungen von Lehrkräften in Bezug auf Leistungsbeurteilungen (Hesse & Latzko, 2017; Stang, 2016) und moderate bis starke Korrelationen (Lorenz, 2011) werden berichtet. Da sich die bisherigen Forschungsergebnisse meist auf den Sekundarschulbereich oder auf andere Teilleistungsstörungen beziehen (Karing. 2009), stellt sich die Frage, wie gut Grund- und Förderschullehrkräfte eine Rechenstörung in der dritten Klasse identifizieren können.

Methode:

An der Studie nahmen bislang N = 129 Schüler:innen und N = 10 Lehrkräfte aus Grund- und Förderschulen teil. Die Lehrkräfte schätzten die Mathematikleistung der Schüler:innen auf einer Skala von 0 (sehr schlecht) bis 10 (sehr gut) ein. Zusätzlich beantworteten die Schüler:innen vier Likert-skalierte Items zur Motivation und Selbsteinschätzung in Mathematik. Die tatsächliche Mathematikleistung wurde mit dem Heidelberger Rechentest (HRT) (Haffner et al., 2005) erfasst. Zur Überprüfung der Zusammenhänge zwischen den Lehrerurteilen und der tatsächlichen Mathematikleistung wurden Korrelationskoeffizienten nach Spearman berechnet. Anschließend wurden die Koeffizienten mit der Fisher's z-Transformation verglichen, um einen statistischen Unterschied zwischen der Beurteilung von Grund- und Förderschullehrkräften zu untersuchen. Zusätzlich wurde ein multiples Regressionsmodell zur Analyse von Einflüssen auf die Mathematikleistung angepasst. Die Berechnung der Güte der Lehrkrafturteile erfolgte anhand des Kriteriums der Sensitivität getrennt für beide Schultypen.

Ergebnisse:

Die Analysen ergaben erwartungskonform mit den Ergebnissen von Lorenz (2011) einen signifikanten Zusammenhang zwischen der Einschätzung der Lehrkräfte und der tatsächlichen Mathematikleistung (r = .74, p < .001). Dieser Zusammenhang ist nur in der Grundschule signifikant (r = .68, p < .001; r = .25, p > .05). Der anschließende Vergleich der Korrelationskoeffizienten zeigte einen signifikanten

Unterschied (z = 2.4, p < .05). Scheinbar ist die Urteilsgüte zwischen Grund- und Förderschulen signifikant voneinander verschieden. Ebenso zeigte sich ein Unterschied in der Sensitivität der Urteile: Die Urteilsgüte ist bei den Förderschullehrkräften mit einer Sensitivität von 70 % besser als die der Grundschullehrkräfte. Nur in 25 % der Fälle gelingt es Grundschullehrkräften eine Rechenstörung richtig zu identifizierten.

Weiterhin schien die Leistungsbeurteilung der Lehrkräfte, die Selbsteinschätzung sowie Motivation und Schulform einen signifikanten Einfluss auf die mathematische Testleistung zu haben (F(4, 124) = 66.19, p = <.001, R^2 = 0.67).

Die Stichprobengröße soll in einer im Herbst 2022 geplanten Erhebung an weiteren Grund- und Förderschulen erweitert werden, um Mehrebenenanalysen anzupassen und den Unterschied in der Urteilsgüte zwischen Grund- und Förderschullehrkräften anhand einer größeren Stichprobe zu schätzen. Bereits die vorliegenden Ergebnisse deuten auf einen Unterschied in der Urteilsgüte von Grund- und Förderschullehrkräften hin. Ebenso weisen die Ergebnisse auf einen Mangel an diagnostischer Kompetenz bei Grundschullehrkräften hin, welche als Teil des Professionswissen und als Voraussetzung für eine gute Unterrichtsqualität daher unbedingt gefördert werden sollte.

Urteilsverzerrung bei der Diagnose von Fehlkonzepten bei Dezimalbrüchen

Andreas Rieu, Timo Leuders, Katharina Loibl
PH Freiburg, Germany

Theoretischer Hintergrund

Fehlkonzepte von Lernenden stellen einerseits eine besondere Herausforderung für Lehrkräfte dar, da sie Lernprozesse von Lernenden behindern können. Weil Fehlkonzepte aber auch das Schülerdenken abbilden, bieten sie andererseits eine wichtige Gelegenheit für die Lehrkraft, adaptive unterrichtliche Maßnahmen einzuleiten. Bei manchen Lerngegenständen – wie beispielsweise bei Dezimalzahlen – führen verschiedene Fehlkonzepte zum selben, fehlerhaften Schülerverhalten (Stacey, 2005) und eine möglichst eindeutige und akkurate Diagnose des zugrunde liegenden Fehlkonzeptes ist daher von entscheidender Bedeutung (Beck et al., 2008).

Auch wenn Urteile in pädagogischen Situationen und ihre Akkuratheit seit geraumer Zeit stärker ins Forschungsinteresse gerückt sind (Heitzmann et al., 2019; Loibl et al., 2020; Südkamp et al., 2012), so mangelt es dennoch an belastbaren Erklärungen für die interindividuell stark variierende Urteilsakkuratheit von Lehrkräften bei der Einschätzung von Lernenden. Die Modellierung des Lehrkräftedenkens bei der Diagnose von Fehlkonzepten in Anlehnung an den Prozess des sozialen Hypothesentestens (Trope & Liberman, 1996) lässt mögliche Ursachen hierfür identifizieren: Aufgrund der Mehrdeutigkeit der Diagnose von Fehlkonzepten bei Dezimalbrüchen müssen diagnostisch relevante Information verarbeitet werden, wobei kognitive Verzerrungen (sog. *biases*) auftreten können (Oswald & Grosjean, 2004). Konkret kann 1.) die Mehrdeutigkeit der Situation unberücksichtigt bleiben (Formulierung einer Einfach- anstelle einer Mehrfachhypothese) oder 2.) die nachfolgende Informationsverarbeitung nur im Sinne der Bestätigung eines angenommen Fehlkonzeptes durchgeführt werden (Auswahl weiterer, das angenommene Fehlkonzept bestätigende Aufgaben und Interpretation ihrer Diagnoserelevanz). Es ist anzunehmen, dass beide

Urteilsverzerrungen a) auftreten und b) durch gezielte Interventionen reduziert werden können. Empirische Evidenz im Bereich diagnostischer Urteile von Lehrpersonen liegen hierzu bislang allerdings nicht vor.

Fragestellung

In der vorliegenden Studie wird untersucht, ob bei der mehrdeutigen Situation der Diagnose von Fehlkonzepten beim Dezimalbruchvergleich kognitive Verzerrungen auftreten. Aufgrund der modellierten Urteilssituation wird folgende Forschungsfrage abgeleitet:

Lässt sich die Urteilsakkuratheit durch 1.) die Formulierung einer Mehrfachhypothese bei der Ersthypothese und 2.) die Auswahl von diagnostisch relevanten Aufgaben bei der Informationssuche vorhersagen?

In anderen Bereichen konnte gezeigt werden, dass kognitive Verzerrungen durch spezifische Interventionen reduziert werden können (Larrick, 2004). Vor diesem Hintergrund wird angenommen, dass gezielte Interventionen auch bei der Diagnose von Fehlkonzepten zur Reduktion von kognitiven Verzerrungen führen. Untersucht wird:

Welchen Einfluss haben Interventionen bezüglich 1.) der Wahrnehmung der Mehrdeutigkeit der Urteilssituation und 2.) der Informationsverarbeitung auf die Urteilsakkuratheit?

Methode

Zur Überprüfung dieser theoretischen Annahmen wurden angehende Lehrkräfte mit jeweils einer fehlerhaften Schülerlösung aus dem Bereich Dezimalbruchvergleiche konfrontiert und aufgefordert, Diagnosehypothesen zu stellen. Anschließend war das Ziel die eindeutige Diagnose des vorliegenden Fehlkonzepts. Dazu konnten die teilnehmenden Personen weitere Aufgaben auswählen, die von den zu diagnostizierenden Lernenden gelöst wurden. Diese Aufgaben unterschieden sich in ihrem diagnostischen Gehalt: Sie waren entweder relevant oder irrelevant zur Diagnose des vorliegenden Fehlkonzepts.

Zur Beantwortung der Forschungsfrage wurden die Urteilsprozesse in fünf Diagnosen von angehenden Lehrkräften (N = 79) untersucht. Erfasst wurde 1.) die Art der erstellten Ersthypothesen (Einfach- oder Mehrfachhypothesen), 2.) die Anzahl und die Art (diagnostisch relevant oder irrelevant) der verarbeiteten Informationen in Form zusätzlicher Aufgaben und 3.) die Akkuratheit der Enddiagnose. Zur Schätzung der Effekte der Prädiktoren und aufgrund der erwartbaren Abhängigkeiten innerhalb der Daten (Personen- oder Testeffekte) wurden verallgemeinerte lineare gemischte Modelle mit der Urteilsakkuratheit als abhängige Variable verwendet/berechnet.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass der Anteil der verwendeten relevanten Aufgaben einen signifikanten Effekt auf die Urteilsakkuratheit hat (OR = 2.90, 95% CI [1.38 – 6.09], p = .005), die Formulierung einer Mehrfachhypothese hingegen erhöht nur deskriptiv, aber nicht signifikant, die Wahrscheinlichkeit einer korrekten Diagnose (OR = 1.70, 95% CI [0.87 – 3.38], p = .120). Die Auswertung der Effekte der Interventionen steht noch aus.

Die Ergebnisse liefern Impulse für die weitere Forschung zum diagnostischen Denken und Handeln sowie für die Lehrkräfteausbildung im Bereich diagnostischer Kompetenzen.

B-05: Grundschule

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S06 S03 B06
Chair der Sitzung: Nadine Schuler

Validierung eines Messinstruments zur Erfassung von selbstreguliertem Lernen auf Grundschulstufe

Nadine Schuler, Caroline Villiger, Evelyn Krauß
PH Bern, Switzerland

Theoretischer Hintergrund

In einer sich rasch verändernden Welt, in der laufend neue Anforderungen an den Menschen gestellt werden, gewinnt das selbstregulierte Lernen (SRL) zunehmend an Bedeutung. Bei dem SRL steht das Individuum und sein aktiver, konstruktiver Lernprozess im Zentrum (Otto et al., 2015). Selbstreguliert Lernende verfügen zumeist über mehr Strategiewissen und zeigen bessere Leistungen (Karlen, 2016). Jedoch zeigen sich beim SRL Unterschiede hinsichtlich des Geschlechts: Mädchen geben im Vergleich zu Jungen häufiger an selbstreguliert zu lernen (Vandevelde et al., 2013; Zuffiano et al., 2012).

Das SRL gliedert sich nach Zimmerman (1990) in einem zyklischen Modell in die drei Phasen «Planung», «Handlung» und «Reflexion» und basiert auf dem sozial-kognitiven Ansatz nach Bandura (1986). Ausgehend von Zimmermans Theorie haben Stöger und Ziegler (2008) ein Prozessmodell mit sieben Phasen (Selbsteinschätzung, Zielsetzung, Planung, Strategieanwendung, Strategiemonitoring, Strategieanpassung und Bewertung des Ergebnisses) ausdifferenziert. Dieses Prozessmodell ermöglicht eine nuancierte Darstellung der verschiedenen Phasen des SRL. Benick et al. (2021) sehen eine Differenzierung der Planungsphase in «Strategieplanung» und «Zeitplanung» vor. Diese Unterscheidung ist naheliegend, da bei der Planung nicht nur strategische, sondern auch zeitliche Aspekte bedeutsam sind.

Es bestehen zahlreiche Messinstrumente zur Erfassung von SRL. Bei diesen handelt es sich zumeist um fachspezifische Erhebungsinstrumente (z.B. Otto, 2007), Instrumente zur Typisierung (z.B. Ziegler et al., 2010) oder Instrumente mit Fokus auf Lernstrategien (z.B. González-Torres & Torrano, 2008). Weiter existieren Instrumente zur Erfassung des SRL auf gymnasialer und universitärer Stufe, welche aufgrund ihrer Komplexität nicht auf der Grundschulstufe eingesetzt werden können (z.B. Karlen et al., 2018, Pintrich et al., 1991). Benick und Mitarbeitende (2021) konzipierten ein Messinstrument zur Erhebung des SRL auf Grundschulstufe. Dieses erfasst jedoch nicht alle Phasen des Prozessmodells von Stöger und Ziegler (2008). Ein geeignetes Instrument zur Erfassung von SRL bei Grundschüler/-innen, welches sämtliche Phasen des selbstregulierten Lernens abbildet, liegt derzeit nicht vor.

Fragestellung

Im Zentrum dieses Beitrags steht die Erfassung des SRL auf der Grundschulstufe mittels eines Fragebogens. Folgende Forschungsfragen werden beantwortet:

- Kann das SRL empirisch in einem Modell zweiter Ordnung (anhand der 8 Faktoren) abgebildet werden? Ist die Reliabilität befriedigend?

- Zeigen Lernende, die häufiger selbstreguliert lernen, auch höhere Leistungen und mehr Strategiewissen? (Prognostische Validität)

- Unterscheidet sich das SRL der Schüler/-innen hinsichtlich des Geschlechts (Differenzielle Effekte)?

Methode

Die Daten wurden im Herbst 2021 innerhalb einer Interventionsstudie zur Förderung des SRL erhoben. Befragt wurden N = 771 Schüler/-innen aus 40 Klassen der 5./6. Klassenstufe (Alter t1 M= 11.28 Jahre, SD = 0.70, weiblich = 51.6%).

Die Struktur des SRL-Modells (Selbsteinschätzung, Zielsetzung, Zeitplanung, Strategieplanung, Strategieanwendung, Strategiemonitoring, Strategieanpassung und Bewertung) wurde anhand einer latenten konfirmatorischen Faktorenanalyse mittels Mplus (Muthén & Muthén, 2004) getestet. Jeder der acht Faktoren wurde anhand von drei Items erfasst. Nach der Prüfung der Fit-Indices wurde die prognostische Validität in Bezug auf Leseleistung und Strategiewissen betrachtet. Weiter wurden differenzielle Effekte hinsichtlich des Geschlechts analysiert.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse zeigen ausreichende bis gute Fit-Indices. Zudem überzeugt die Erfassung des SRL mittels acht Faktoren im Vergleich zu einer Drei-Faktoren- (Zimmerman) bzw. Sieben-Faktoren-Lösung (Stöger & Ziegler, 2008) mit dem besten Modellfit.

Die prognostische Validität des Erhebungsinstruments kann aufgrund erster Befunde insofern als zufriedenstellend erachtet werden, als das sich Zusammenhänge zwischen dem SRL und dem Strategiewissen ergeben. Ebenfalls lassen sich mit dem Messinstrument frühere Forschungsergebnisse zu differenziellen Effekten weitgehend replizieren.

Ein passendes Messinstrument zur Erfassung des SRL auf Grundschulstufe lag bisher nicht vor. Das entwickelte Messinstrument bietet einen differenzierten, schulfachunabhängigen Fragebogen zur Erfassung des SRL und eröffnet damit einen neuen Ansatz zur Messung des SRL auf Grundschulstufe.

Indirekte Effekte der Metakognition bei Grundschüler:innen auf die Schulempfehlungen und Übertrittsentscheidung

Sog Yee Mok[1,2], Kathrin Lockl[3], Markus P. Neuenschwander[2]
[1]Pädagogische Hochschule Graubünden, Schweiz; [2]Pädagogische Hochschule FHNW, Schweiz; [3]Leibniz-Institut für Bildungsverläufe (LIfBi), Bamberg

Theoretischer Hintergrund

Lehrkräfteurteile beeinflussen die Schullaufbahn der Schüler:innen maßgeblich. So wurde beispielsweise gezeigt, dass die Lehrkräfteurteile der mathematischen Fähigkeiten und Lesefähigkeiten bedeutsam für die Empfehlungen der Lehrkräfte für weiterführende Schulen nach der Grundschule sind (Maaz & Nagy, 2010). Jedoch beurteilen Lehrkräfte nicht nur die Leistungen der Schüler:innen, sondern auch andere Fähigkeiten wie die Metakognition, die im Unterricht weniger direkt sichtbar ist. Die Metakognition, insbesondere das deklarative Metakognitionswissen, welches sich insbesondere auf das Wissen über (Lern-)Strategien bezieht (Flavell, 1979), ist ein starker Prädiktor für mathematische Fähigkeiten und das Leseverständnis (Artelt et al., 2012; Desoete, Roeyers & Buysse, 2001). Bislang zeigen nur wenige Studien, dass Lehrkräfteurteile zur Metakognition von Schüler:innen auch einen Einfluss auf Schulempfehlungen für Grundschüler:innen

haben (Boone & Van Houtte, 2013). Die Schulempfehlungen der Lehrkräfte beeinflussen wiederum stark den tatsächlichen Schulübertritt in das Gymnasium (Neuenschwander et al., 2015). Vor diesem Hintergrund hat der vorliegende Beitrag das Ziel, die indirekte Wirkung der Metakognition bei Schüler:innen in der Grundschule über Lehrkräfteurteile der fachspezifischen Fähigkeiten der Schüler:innen auf Schulempfehlungen sowie Schulübertritte zu untersuchen. Da bei den Schulempfehlungen der Lehrkräfte nicht nur die Leistung in Mathematik und das Leseverständnis in Deutsch, sondern auch weitere Faktoren wie das Geschlecht, der sozioökonomische Status und die kognitiven Grundfähigkeiten eine Rolle spielen, werden diese Aspekte in den Analysen kontrolliert.

Fragestellungen

Frage 1: Hat die Metakognition von Grundschüler:innen eine indirekte Wirkung auf die Schulempfehlungen, welche durch die Lehrkräfteurteile zu mathematischen und schriftsprachlichen Fähigkeiten vermittelt wird?

Frage 2: Hat die Metakognition von Grundschüler:innen eine indirekte Wirkung auf die tatsächliche Schulübertrittsentscheidung, welche durch die Lehrkräfteurteile zu mathematischen und schriftsprachlichen Fähigkeiten und Schulempfehlungen vermittelt wird? Mögliche Effekte der Mathematikleistung, des Leseverständnisses, des Geschlechts, des sozioökonomischen Status und der kognitiven Grundfähigkeiten werden dabei kontrolliert.

Methode

Stichprobe

In der Studie wurden Daten aus der Kindergartenkohorte des Nationalen Bildungspanels Deutschland (NEPS) verwendet. In der Stichprobe waren 5870 Kinder (51,4 % Mädchen) aus 890 verschiedene Grundschulen. Das Durchschnittsalter der Kinder in Klasse 4 betrug 9 Jahre und 9 Monate (*SD* = 4,6 Monate).

Instrumente

Metakognition wurde mittels deklarativen Metakognitionswissen mit einem Szenario-basierten Kompetenztest (10 Szenarios) in der 3. Klasse erhoben (Artelt, Beinicke, Schlagmüller, & Schneider, 2009; Lockl, Händel, & Artelt, 2018). Dieser Kompetenztest fragte das Wissen über die Angemessenheit von verschiedenen (Lern-)Strategien in unterschiedlichen Situationen ab. Die Fähigkeiten in Mathematik und im Leseverständnis sowie die Lehrkräfteurteile zu mathematischen und schriftsprachlichen Fähigkeiten (Lesen und Schreiben; die Beurteilungen erfolgten pro Kind; Skala von 1 bis 5) wurden im 4. Schuljahr erfasst. Die Schulempfehlungen der Lehrkräfte, sowie die tatsächliche Schulübertrittsentscheidung wurden dichotom erfasst (1 = Gymnasium vs. 0 = andere Schultypen).

Analysen

Es wurde ein Mehrebenen-Pfadmodell in Mplus 7.4 gerechnet. Dieses beinhaltete direkte Pfade der Metakognition, der Testergebnisse in den Bereichen Mathematik und Lesen, von HISEI, Geschlecht und kognitiven Grundfähigkeiten auf die Lehrkräfteurteile der mathematischen und schriftsprachlichen Fähigkeiten, auf die Schulempfehlung der Lehrkräfte sowie auf den Gymnasiumsübertritt. Weiterhin

wurden direkte Pfade von den Lehrkräfteurteilen auf die Schulempfehlung und den Schulübertritt sowie von der Schulempfehlung auf den Gymnasiumsübertritt spezifiziert. Die FIML-Funktion in Mplus wurde für die fehlenden Werte eingesetzt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigten einen signifikanten indirekten Effekt der Metakognition auf die Gymnasiumsempfehlungen von Lehrkräften, die durch die Lehrkräfteurteile zu mathematischen Fähigkeiten und schriftsprachlichen Fähigkeiten vermittelt wurden. Weiter zeigte sich ein signifikanter indirekter Effekt der Metakognition auf den Gymnasiumsübertritt nach der Grundschule, der durch die Lehrkräfteurteile und Schulempfehlungen von Lehrkräften vermittelt wurde. Die indirekten Effekte der Metakognition auf die Schulempfehlungen und den Gymnasiumsübertritt verdeutlichen die Bedeutsamkeit der Metakognition bei Grundschüler:innen, für die Leistungsbeurteilung und Übertrittsentscheidungen ins Gymnasium. Obwohl die Metakognition kein Übertrittskriterium ist, fließt es in die Übertrittsentscheidungen ein. Dies sollte in der Lehrkräfteausbildung diskutiert werden.

Subjektives Wohlbefinden und akademische Leistung in der Grundschule: Längschnittliche Befunde von der 3. bis zum 5. Klasse?

Ineke Pit-ten Cate, Pascale Esch, Ulrich Keller, Antoine Fischbach, Rachel Wollschläger
Luxembourg Centre for Educational Testing, University of Luxembourg, Luxembourg

Sowohl schulische Leistungen als auch subjektives Wohlbefinden sind integrale Ziele (inklusiver) Bildung (Hascher et al., 2018). Subjektives Wohlbefinden im Bildungskontext umfasst die Freude am Schulbesuch, positive Beziehungen zu Lehrkräften und Mitschüler*innen sowie das schulische Selbstkonzept. Bildungssysteme müssen sowohl die akademischen als auch die sozialen Bedürfnisse der Schüler berücksichtigen und sicherstellen, dass alle Schüler als geschätzte und integrale Mitglieder der Schulgemeinschaft aktiv an den Aktivitäten teilnehmen können. Die Forschung hat jedoch gezeigt, dass die soziale Teilhabe durch demografische Merkmale und schulische Leistungen von Schülern beeinflusst werden kann, wodurch Schüler mit Migrationshintergrund oder sonderpädagogischem Förderbedarf größere Schwierigkeiten haben, positive soziale Beziehungen aufzubauen und aufrechtzuerhalten (Matheson et al., 2007; OECD, 2017) und leistungsschwache Schüler sind einem erhöhten Risiko ausgesetzt, gemobbt zu werden (OECD, 2019).

In dieser Studie haben wir den Zusammenhang zwischen subjektivem Wohlbefinden und akademischer Leistung und sozio-demografischer Variablen untersucht. Ein erstes Ziel war es, diesen Zusammenhang in der 3. Klasse zu prüfen (Querschnittdaten). Ein weiteres Ziel der Studie bestand darin, den Zusammenhang zwischen subjektivem Wohlbefinden und Schülermerkmalen (Leistung, Geschlecht und Migrationshintergrund) zu ermitteln.

Die Ergebnisse basieren auf einer Kohorte von Schüler*innen, die von der 3. bis zur 5. Klasse am nationalen Schulmonitoring (Luxembourg School Monitoring Programm „Épreuves Standardisées“; (Martin et al., 2015)) teilgenommen haben (*N*=3733). Die Daten im November 2019 (3. Klasse) und November 2021 (5. Klasse) in Luxemburg erhoben. Im Rahmen dieser Erhebung wurden sowohl standardisierte Schulleistungstests als auch ein Fragebogen zu soziodemographischen und sozio-emotionalen Aspekten durchgeführt. Indikatoren für subjektives Wohlbefinden umfassen allgemeines akademisches Selbstkonzept, soziale und emotionale Inklusion (alle gemessen mit dem Fragebogen

zur Wahrnehmung der Inklusion, PIQ (Venetz et al., 2014). Für die schulischen Leistungen verwenden wir ÉpStan-Ergebnisse für Deutschleseverständnis und Mathematik. Zusätzlich wurden über einen Schüler- und Elternfragebogen weitere sozio-demographische Merkmale erfasst.

In einer ersten (Querschnitt) Regressionsanalyse haben wir die Domäne des subjektiven Wohlbefindens der Schüler*innen in der 3. Klasse anhand individueller Schülermerkmale (Leistung, Geschlecht und Migrationshintergrund) vorhergesagt. Zweitens und unter Verwendung der Längsschnittdaten, wurde die Domäne des subjektiven Wohlbefindens der Schüler*innen in der 5. Klasse über eine schrittweise Regression unter statistischer Berücksichtigung des Wohlbefindens in der 3. Klasse vorhergesagt: Zunächst wurden sozio-demographische Hintergrundvariablen (Geschlecht und Migrationshintergrund) dem Model hinzugefügt und, in einem zweiten Schritt, Indikatoren der akademischen Leistung.

Die ersten Ergebnisse zeigten einen Zusammenhang zwischen dem subjektiven Wohlbefinden in der 3. Klasse, dem Geschlecht, dem Migrationshintergrund der Schüler*innen und ihrer Leistung. Dieser Zusammenhang ist stärker für das akademische Selbstkonzept als für die soziale und emotionale Inklusion. Weiterhin zeigten die Ergebnisse basierend auf den Längsschnittdaten, dass unter statistischer Berücksichtigung des subjektiven Wohlbefindens in der 3. Klasse, nur bis zu 1% der Varianz des subjektiven Wohlbefindens in der 5. Klasse durch soziodemografische Variablen (insbes. Geschlecht) vorausgesagt werden kann. Die akademische Leistung erklärte zusätzliche 1% bis 11% der Varianz in der Wahrnehmung der sozialen und emotionalen Inklusion bzw. dem akademische Selbstkonzept.

Die Ergebnisse dieser Studie zeigten, dass das Wohlbefinden der Schüler*innen bereits in der 3. Klasse etabliert ist und im Zeitverlauf nur geringfügig von soziodemografischen und Leistungsvariablen beeinflusst wird. Ein wichtiger Aspekt ist, dass die erste Erhebung vor der Covid 19-Pandemie erfolgt ist, während die zweite Erhebung während der Pandemie, allerdings nicht während einer Phase der Schulschließungen, erfolgte. Vor diesem Hintergrund scheint das relative stabile Wohlbefinden besonders nennenswert zu sein. Limitierend ist zu erwähnen, dass nur Schüler*innen in der Studie berücksichtigt werden konnten, die eine reguläre Schullaufbahn hatten und keine Klasse wiederholten.

B-06: Diagnostik durch Lehrkräfte

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S06 S04 B06
Chair der Sitzung: Katharina Schnitzler

„Diese Schülerin ist im Unterricht sehr motiviert!“ Wie soziodemographische Merkmale von Schüler:innen die Einschätzungen und Notenvergabe von Lehrpersonen beeinflussen

Katharina Schnitzler[1], Cornelius Brandmiller[2,3], Hanna Dumont[1]
[1]Universität Potsdam; [2]DIPF, Leibniz-Institut für Bildungsforschung und Bildungsinformation; [3]Senatsverwaltung für Bildung, Jugend und Familie, Berlin

Theoretischer Hintergrund

Noten sind als die gängigste Form der Leistungsbewertung integraler Bestandteil von Schule (Brookhart, 1994). Aufgrund ihrer gesellschaftlichen und pädagogischen Funktionen beeinflussen sie Bildungswege maßgeblich, indem Schüler:innen beispielsweise erst ab einem bestimmten Notenschnitt Zugang zu gewissen Schularten und Studiengängen gewährt wird (Elsäßer, 2018). Lehrpersonen tragen deshalb bei der Notenvergabe eine große Verantwortung; diese sollte maßgeblich auf akkuraten Einschätzungen der fachlichen Leistung basieren. Studien zeigen jedoch, dass Lehrpersonen auch ihre Einschätzungen bezüglich motivationaler Merkmale von Schüler:innen in Noten abbilden (Brookhart et al., 2016). Dabei ist problematisch, dass Lehrpersonen insbesondere bei der akkuraten Einschätzung der Motivation Schwierigkeiten zeigen und soziodemographischen Merkmale von Schüler:innen in die Urteile einfließen (Urhahne & Wijnia, 2021; Wang et al., 2018). So scheinen Lehrpersonen die Motivation von Mädchen, von Schüler:innen mit höherem sozioökonomischen Status (SES) und von Schüler:innen, die keinen oder einen osteuropäischen Migrationshintergrund haben zu überschätzen (Brandmiller et al., 2020; Gentrup et al., 2018). Da sich ähnliche Effekte auch für Schulnoten finden (z.B. Autin et al., 2019), wäre es denkbar, dass diese durch die beschriebenen Unterschiede in den motivationalen Einschätzungen von Lehrpersonen erklärt werden können.

Fragestellung

Ziel des Beitrags ist es, die Rolle der motivationalen Einschätzungen von Lehrpersonen im Rahmen der Notenvergabe in Deutsch und Mathematik zu untersuchen:

1. Beeinflussen soziodemographische Merkmale von Schüler:innen die motivationalen Einschätzungen von Lehrpersonen zu Beginn der 5. Jahrgangsstufe?
2. Stellen motivationale Einschätzungen von Lehrpersonen einen Mediator zwischen soziodemographischen Merkmalen und Schulnoten in der 5., 6., und 7. Jahrgangsstufe dar?

Methode

Datengrundlage war eine Teilstichprobe von 3400 Schüler:innen sowie deren 58 Deutsch- und 50 Mathematiklehrpersonen einer großen Schulleistungsstudie.

Die Lehrpersonen schätzten die Motivation ihrer Schüler:innen zu Beginn der 5. Jahrgangsstufe ein (5 Items, α = .91, z.B. Diese:r Schüler:in ist im Unterricht sehr motiviert). Für beide Fächer wurden die Schuljahresabschlussnoten der 5., 6., und 7. Jahrgangsstufe erfragt. An soziodemographischen Variablen wurden SES, Geschlecht und Migrationshintergrund erfasst während die Noten der 4. Jahrgangsstufe, standardisierte Leistungen und selbstberichtete Motivation der Schüler:innen als Kontrollvariablen berücksichtigt wurden.

Mit Strukturgleichungsmodellen wurden die Daten für Deutsch und Mathematik getrennt analysiert. Für die erste Forschungsfrage wurden die Einschätzungen der Lehrpersonen durch die

soziodemographischen Merkmale der Schüler:innen vorhergesagt. Für die zweite Forschungsfrage wurden direkte und indirekte Effekte (vermittelt über die Lehrpersoneneinschätzungen) von den demographischen Merkmalen auf die Noten der 5. bis 7. Jahrgangsstufe geschätzt.

Ergebnisse und Bedeutung

Bezüglich der ersten Forschungsfrage zeigte sich, dass Lehrkräfte die Motivation von Jungen in Deutsch (β = -.34, $p < .001$) und Mathematik (β = -.62, $p < .001$) signifikant niedriger einschätzten als die von Mädchen. Außerdem schätzten Deutschlehrpersonen die Motivation von Schüler:innen mit Migrationshintergrund signifikant niedriger ein als die von Schüler:innen ohne Migrationshintergrund (β = -.22, $p = .027$) während Mathematiklehrpersonen diesbezüglich keinen Unterschied machten. Für den SES der Schüler:innen ergaben sich keine signifikanten Zusammenhänge.

Hinsichtlich der zweiten Fragestellung zeigte sich, dass Jungen in Deutsch und Mathematik von der 5. bis zur 7. Jahrgangsstufe vermittelt über die motivationalen Einschätzungen der Lehrkräfte schlechtere Noten erhielten als Mädchen (β zwischen -.22 und -.11, alle $p < .05$) und, dass Schüler:innen mit Migrationshintergrund in der 5. Jahrgangsstufe in Deutsch schlechtere Noten bekamen als Schüler:innen ohne Migrationshintergrund (β = -.07, $SE = .03$, $p = .027$). Alle anderen indirekten Effekte waren nicht signifikant.

Die Ergebnisse verdeutlichen, dass die motivationalen Einschätzungen eine Rolle dabei spielen wie Lehrpersonen ihre Schüler:innen benoten. Wobei die Einschätzungen scheinbar überwiegend zu geschlechtsspezifischen Unterschieden in den Noten beitragen.

Diagnostische Kompetenz von Lehrpersonen im selbstregulierten Lernen: Wissen und Urteilsgenauigkeit

Kerstin Bäuerlein, Sabrina Brunner, Yves Karlen
Pädagogische Hochschule FHNW, Switzerland

Kompetenzen im selbstregulierten Lernen (SRL) sind zentrale Bildungsziele (OECD, 2019) und gehen mit grösserem Erfolg in Schule und Beruf einher (Dent & Koenka, 2016; Rademacher, 2019; Schneider & Preckel, 2017). Selbstregulierte Schüler:innen planen, beobachten und regulieren ihren Lernprozess und übernehmen zunehmend Verantwortung dafür (Pressley et al., 1989). Viele Schüler:innen jeden Alters haben jedoch Mühe mit dem SRL (Karlen, 2016; McCabe, 2011). Daher ist die Diagnose und Förderung des SRL wichtig. Studien weisen jedoch darauf hin, dass Lehrpersonen die Förderung von SRL bislang wenig in ihren Unterricht integrieren (Dignath & Veenman, 2021).

Für die Förderung von SRL benötigen Lehrpersonen verschiedene professionelle Kompetenzen (Karlen et al., 2020): Ein tiefes Verständnis von SRL und Methoden zur SRL-Diagnose sowie die Fähigkeit, die Diagnosemethoden anzuwenden und den Unterricht an den diagnostizierten Kompetenzstand der Schüler:innen anzupassen. Bisher wurde das Verständnis von Lehrpersonen zum SRL und dessen Diagnose wenig untersucht (Dignath & Sprenger, 2020; Michalsky, 2017). Auch die Frage, wie akkurat Lehrpersonen das SRL ihrer Schüler:innen einschätzen können, ist unzureichend geklärt. Vorliegende Studien weisen auf eine geringere Akkuratheit der Lehrpersonenurteile bei nicht-kognitiven Merkmalen im Vergleich zu kognitiven Merkmalen hin (Lee &

Reeve, 2012; Spinath, 2005; Urhahne & Wijnia, 2021). Weitere Erkenntnisse hierzu sind hochbedeutsam, da eine adäquate Diagnose Ausgangspunkt für effektive Fördermassnahmen ist.

Hier setzt die vorliegende Studie an, indem sie folgende Fragestellungen untersucht:

1. Was wissen Lehrpersonen über SRL und dessen Diagnose?
2. Wie akkurat schätzen Lehrpersonen das SRL ihrer Schüler:innen ein?
3. Bestehen Zusammenhänge zwischen dem Wissen der Lehrpersonen und ihrer Urteilsgenauigkeit?

Um diese Fragestellungen zu beantworten, wurden N = 51 Lehrpersonen der Sekundarstufe I (41.5% weiblich, mittlere Berufserfahrung 15 Jahre, SD = 8.2) von vier Schweizer Schulen rekrutiert. Die Lehrpersonen und ihre Klassen nahmen freiwillig an der Studie teil. Die Lehrpersonen bearbeiteten einen Online-Fragebogen, in dem sie demographische Angaben machten und offene Fragen zu ihrem Verständnis von SRL und dessen Diagnose beantworteten. Zudem schätzten sie das Wissen der ersten fünf Schüler:innen ihrer Klassenliste zu vier SRL-Kompetenzbereichen (Monitoring und Reflexion, Management der Lernumgebung, Selbstkontrolle sowie Motivationsregulation) auf einer vierstufigen Skala ein. Von diesen Schüler:innen nahmen N = 173 an der Studie teil (49.7% weiblich, mittleres Alter 14.42 Jahre, SD = 0.97). Sie bearbeiteten Wissenstests zu den vier genannten SRL-Kompetenzbereichen ($.60 \leq \alpha \leq .88$).

Die offenen Antworten wurden zunächst mittels strukturierender Inhaltsanalyse ausgewertet (Mayring, 2015) und anschliessend mittels evaluativer Inhaltsanalyse (Kuckartz & Rädiker, 2022) auf einer vierstufigen Ordinalskala eingeordnet. Alle Antworten wurden durch zwei Projektmitarbeiter:innen konsensuell kodiert (Schmidt, 2013). Die Urteilsgenauigkeit der Lehrpersonen bezüglich des SRL ihrer Schüler:innen wurde mittels Korrelationen (unter Berücksichtigung der Clusterung) zwischen Lehrpersoneneinschätzung und Schüler:innenergebnissen im Wissenstest bestimmt.

Die Ergebnisse zeigen eine grosse Variabilität des Wissens von Lehrpersonen über SRL und dessen Diagnose, wobei in Einklang mit Vorbefunden (Dignath & Sprenger, 2020; Glogger-Frey et al., 2018; Ohst et al., 2015) bei knapp 20 % der Lehrpersonen Fehlkonzepte evident wurden. Hinsichtlich der von den Lehrpersonen gemäss Selbstbericht eingesetzten Methoden zur Erfassung von SRL wurde konform mit früheren Studienergebnissen (Michalsky, 2017) deutlich, dass Lehrpersonen eher Offline-Methoden (vor oder nach der Aufgabenbearbeitung) als Online-Methoden (während der Aufgabenbearbeitung) verwenden. Die Urteilsgenauigkeit der Lehrpersonen bezüglich des SRL-Wissens ihrer Schüler:innen fiel wie in vorherigen Untersuchungen (Lee & Reeve, 2012; Spinath, 2005; Urhahne & Wijnia, 2021) gering aus ($r \leq .19$, $p \leq .040$). Zudem korreliert das Wissen der Lehrpersonen über SRL in Übereinstimmung mit anderen Studienergebnissen (z. B. Kramer et al., 2021; Michalsky et al., 2017) nicht mit der Urteilsgenauigkeit. Letzteres gilt, wie die vorliegende Studie zeigt, auch für das diagnostische Wissen der Lehrpersonen bezüglich SRL. Künftige Weiterbildungsprogramme sollten daher nicht nur das Wissen der Lehrpersonen über SRL und dessen Diagnose fördern, sondern auch die praktische Anwendung geeigneter Diagnosemethoden.

Profile diagnostischer Urteile von Lehramtsstudierenden und die Rolle des Migrationshintergrunds der Schüler*innen im Fach Mathematik - Ergebnisse einer Interventionsstudie

Christin Laschke[1], Bettina Rösken-Winter[2], Lars Jenßen[2]
[1]Leibniz Institute for Science and Mathematics Education; [2]Humboldt-Universität zu Berlin

Theoretischer Hintergrund

Lehrkräfte sind regelmäßig gefordert, Schüler*innenprodukte zu beurteilen und daraus Schlüsse auf deren Lernstand zu ziehen. Die dabei zum Tragen kommenden diagnostischen Urteile bilden die Basis für die Rückmeldung sowie für daraus resultierende Lernangebote und nehmen somit langfristig Einfluss auf die Entwicklung von Schüler*innen (Zhu et al., 2018). Lehrkräfte hinsichtlich ihrer Urteilsgenauigkeit zu professionalisieren ist daher eine Kernaufgabe des Lehramtsstudiums (Ready & Wright, 2011).

Empirischen Befunden zufolge können diagnostische Urteile durch implizite Stereotype zu bestimmten Merkmalen (Denessen et al., 2022) beeinträchtigt werden, mit der Konsequenz, dass z. B. Schüler*innen mit einem Vornamen, der einen Migrationshintergrund impliziert, bei gleicher Leistung schlechter beurteilt werden als Schüler*innen ohne Migrationshintergrund (Bonefeld et al., 2020).

Bisher gibt es nur wenige Evidenzen dazu, wie dieses Problem mittels Interventionsstudien adressiert werden kann. Erste Befunde dazu zeigen, dass durch eine intensive Auseinandersetzung mit Stereotypen und deren Konsequenzen für Schüler*innen Urteilsverzerrungen aufgrund des Migrationshintergrunds der Schüler*innen im Fach Mathematik reduziert werden können (Neuenschwander et al., 2021).

Weitere Befunde aus der mathematikdidaktischen Forschung deuten darauf hin, dass (angehende) Lehrkräfte über unterschiedliche Urteilsprofile verfügen. Einige Lehrkräfte urteilen vor allem aus fachdidaktischer Perspektive mit Blick auf die Qualität eines Lösungsansatzes, das Verständnis oder die angewandte Prozedur, andere fokussieren auf fachunabhängige Aspekte wie Motivation (Laschke et al., 2022; Baldinger, 2020). Studien zeigen zudem, dass Lerngelegenheiten, die eine intensive Auseinandersetzung mit Schüler*innenprodukten gewährleisten, dazu führen, dass Lehrkräfte weniger fachunabhängige Aspekte und stärker das Verständnis der Schüler*innen fokussieren (Kalinec-Craig et al., 2021). Ungeklärt ist, ob unterschiedliche Urteilsprofile in unterschiedlichem Maße anfällig für Verzerrungen sind und daher unterschiedlicher Interventionen bedürfen.

Vor dem Hintergrund des Problems der Urteilsverzerrungen ist das Ziel der Studie, passgenaue Interventionen für unterschiedliche Gruppen von Lehramtsstudierenden zu entwickeln. Ausgehend von den vorliegenden Evidenzen zur Konzeption von Interventionsstudien wurden folgende Fragestellungen untersucht:

1. Lassen sich unterschiedliche Profile bzgl. diagnostischer Urteile von Lehramtsstudierenden identifizieren?
2. Wird die Zugehörigkeit zu einem bestimmten Profil vor und nach einer Intervention zu fachbezogenen Beurteilungen sowie Stereotypen und deren Konsequenzen für Schüler*innen bedingt durch den Migrationshintergrund der Schüler*innen, deren Produkt beurteilt wird?

Methode

Die Studie wurde in einem randomisierten Experimental-Wartekontrollgruppen-Design mit 209 Primarstufenlehramtsstudierenden am Ende des Bachelorstudiums durchgeführt. Die Experimentalgruppe erhielt eine insgesamt sechsstündige Intervention zu fachbezogenen Beurteilungen anhand von Schüler*innenprodukten sowie zu Stereotypen und deren Konsequenzen für Lernende.

Zu Beginn und am Ende der Studie wurden mathematische Lösungen kombiniert mit unterschiedlichen Vornamen, die laut Vorstudie jeweils (k)einen Migrationshintergrund implizieren. Die Lösungen wurden anhand von zuvor entwickelten Skalen zur Beurteilung des Verständnisses, der Qualität des Lösungsansatzes, der Prozedur und der Motivation (McDonalds ≥ .85) geratet.

Mit den z-standardisierten Beurteilungen der Lehrkräfte wurden Latente Profilanalysen durchgeführt. Nach Prüfung des Vorliegens vollständiger Messinvarianz über beide Erhebungszeitpunkte hinweg, wurde mittels Latent Transition Analyses (LTA) (Hickendorff et al., 2018) untersucht, inwieweit die Zugehörigkeit zu einem Profil abhängig davon ist, ob die Schüler*innenlösung von einem Kind mit Migrationshintergrund stammt.

Ergebnisse

Die Ergebnisse der LTA legen eine Drei-Profile-Lösung nahe (Entropy = .93). In Profil I wurden alle Skalen leicht überdurchschnittlich, in Profil II deutlich überdurchschnittlich geratet, wobei die Skalen *Verständnis* und *Qualität des Lösungsansatzes* besonders hoch geratet wurden. Im Kontrast dazu wurden im Profil III alle Skalen unterdurchschnittlich geratet, die Skalen *Verständnis* und *Qualität des Lösungsansatzes* besonders niedrig.

Den Ergebnissen der LTA zufolge war die Wahrscheinlichkeit zum Profil III zu gehören vor der Intervention in beiden Gruppen sowie nach der Intervention in der Kontrollgruppe höher, wenn die Lösung von einem Kind mit Migrationshintergrund stammte. Jedoch traf dies nicht auf die Experimentalgruppe nach der Intervention zu. Dies lässt vermuten, dass die Intervention den Einfluss impliziter Stereotype minimieren konnte.

Diagnoseprozesse angehender Lehrkräfte in einer digitalen Simulation erklären Unterschiede in der Diagnoseakkuratheit

Christian Schons[1], Andreas Obersteiner[1], Frank Fischer[2], Kristina Reiss[1]
[1]Technische Universität München; [2]Ludwig-Maximilians-Universität München

Theoretischer Hintergrund

Das Diagnostizieren von Lernergebnissen stellt eine Herausforderung für Lehrkräfte dar (Südkamp et al., 2012). Die Förderung von Diagnosekompetenz ist deshalb ein wichtiges Forschungsziel. Dabei umfasst Diagnosekompetenz individuelle Dispositionen (z.B. Professionswissen), situationsspezifische Prozesse und Performanz (Blömeke et al., 2015). Studien deuten zwar darauf hin, dass eine Förderung des Professionswissens (z.B. Rieu et al., 2022) und instruktionale Unterstützungsmaßnahmen während des Diagnoseprozesses (z.B. Sommerhoff et al., 2022) zu besseren Diagnoseergebnissen führen können. Bislang sind aber a) die Diagnoseprozesse selbst und b) deren Zusammenhänge mit Professionswissen einerseits und Diagnoseergebnissen andererseits kaum detailliert beschrieben. Zur Erfassung von Diagnoseprozessen eignen sich Logdaten aus digitalen Simulationen (Goldhammer et al., 2017).

Diese Studie zielt auf die Beschreibung a) individueller Diagnoseprozesse angehender Mathematiklehrkräfte in einer digitalen Simulation und b) der Zusammenhänge individueller Diagnoseprozesse und des Professionswissen mit der Diagnoseakkuratheit. Die Studie verknüpft eine instruktionspsychologische und eine fachdidaktische Perspektive: Die Beschreibung der

Diagnoseprozesse erfolgt auf Basis der drei domänenübergreifenden diagnostischen Aktivitäten *Beschreiben*, *Hypothese generieren*, *Schlussfolgerung ziehen* (Fischer et al., 2014) sowie auf einer fachdidaktischen Kategorisierung der für die Diagnosen ausgewählten Evidenzgrundlage (diagnostisches Potenzial von Mathematikaufgaben, s.u.).

Methode

An der Studie nahmen 65 Studierende des Grundschullehramts (Semesteranzahl M = 2.70, 86% weiblich) im Rahmen von zwei 90-minütigen Seminarsitzungen teil. Zunächst bearbeiteten die Teilnehmenden einen Test (TEDS-M; Tatto et al., 2013) zur Erfassung ihres fachlichen und fachdidaktischen Wissens in Mathematik. Anschließend arbeiteten sie mit einer digitalen Simulation zur Diagnose von Schülerfehlern (Wildgans-Lang et al., 2020). Die Teilnehmenden diagnostizierten darin bis zu acht virtuelle Schülerprofile (Anzahl Bearbeitungen M = 7.45, SD = 1.54) mit dem Ziel, deren mathematische Fehlvorstellungen zu identifizieren. In der Simulation wählten die Teilnehmenden Mathematikaufgaben aus einem Portfolio, legten sie dem virtuellen Schüler bzw. der Schülerin vor und erhielten anschließend eine authentische Schülerlösung zu dieser Aufgabe. Der Vorgang konnte mehrfach wiederholt werden. Die abschließende Diagnose bestand in der Auswahl einer Fehlvorstellung aus einer vorgegebenen Liste.

Die zur Auswahl stehenden Mathematikaufgaben in den Schülerprofilen variierten in ihrem diagnostischen Potenzial (hoch/niedrig), Hinweise auf die spezifischen Fehlvorstellungen der virtuellen Schüler zu geben. Um die drei oben genannten diagnostischen Aktivitäten zu erfassen, erschien zusammen mit der bearbeiteten Schülerlösung ein Textfeld, in welchem die Teilnehmenden ihre Überlegungen notieren sollten. Die Einträge in diesen Textfeldern wurden anschließend nach den drei diagnostischen Aktivitäten kodiert (κ = .84).

Ergebnisse

Eine Clusteranalyse basierend auf den Variablen *Anzahl an gewählten Aufgaben (mit hohem bzw. niedrigem Potential),* sowie den Häufigkeiten der drei diagnostischen Aktivitäten ergab drei Muster an Diagnoseprozessen: 1) Ein Basismuster mit Personen, welche in allen Aktivitäten unterhalb des Mittelwerts lagen, 2) ein Muster an Personen, welche überdurchschnittlich viele Aufgaben (sowohl mit als auch ohne diagnostisches Potential) für ihre Diagnosen auswählten und die Schülerlösungen überwiegend beschrieben, jedoch kaum Hypothesen oder Schlussfolgerungen formulierten, und 3) ein Muster an „effizienten“ Personen, welche wenige Aufgaben für ihre Diagnosen auswählten und überdurchschnittlich viele Hypothesen und Schlussfolgerungen formulierten.

Ein lineares Mischmodell zur Vorhersage der Diagnoseakkuratheit mit fachlichem und fachdidaktischem Wissen als feste Effekte und Schülerprofil sowie Teilnehmende als Zufallseffekte ergab einen signifikanten Einfluss des fachlichen Wissens auf die Diagnoseakkuratheit. Das Professionswissen erklärte einen substantiellen Teil der im Nullmodell unaufgeklärten Varianz der Diagnoseakkuratheit zwischen den Teilnehmenden (PCV = 58.1%; PCV: Nakagawa & Schielzieth, 2019). Hinzufügen der Variable *Prozessmuster* als festen Effekt in das Modell zeigte, dass das effiziente Muster signifikant akkurater diagnostizierte als das Basismuster, und das Gesamtmodell klärte fast die gesamte Varianz zwischen den Teilnehmenden auf (PCV = 98.8%).

Diskussion

Mit Hilfe von Logdaten aus einer digitalen Simulation konnten mathematische Diagnoseprozesse in hoher Auflösung beschrieben werden. Die identifizierten Prozessmuster erklärten Unterschiede in der Diagnoseakkuratheit über das Professionswissen hinaus. Die Ergebnisse können zur Entwicklung adaptiver Prozessunterstützung zur Förderung von Diagnosekompetenz beitragen.

B-07: Sprachsensibilität

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S05 T00 B59
Chair der Sitzung: Laura Beckmann

Sprachsensible Unterrichtspraxis an Schulen in sozialräumlich benachteiligten Lagen. Zur Relevanz des Zuwanderungs- und sprachlichen Hintergrunds der Lehrkräfte

Laura Beckmann, Inka Achtelik
University of Duisburg-Essen, Germany

Theoretischer Hintergrund

Als Folgen der zunehmenden Diversifizierung der Gesellschaft stellt die kulturelle und sprachliche Vielfalt der Schüler:innen im deutschen Schulsystem eine grundlegende Bedingung für (schulische) Bildungsprozesse dar (Brandt, 2021; Gogolin, Klinger, Lagemann & Schnoor, 2017). So wiesen im Schuljahr 2021/22 im Bundesland Nordrhein-Westfalen 40 Prozent der Schüler:innen eine Zuwanderungsgeschichte auf (vgl. IT.NRW, 2022). Nationale und internationale Schulleistungsvergleichsstudien verweisen wiederholt auf erhebliche Leistungsdisparitäten zwischen Schüler:innen mit und ohne Zuwanderungsgeschichte (vgl. Weis et al., 2019). Die jüngste PISA-Studie bestätigt, dass in Deutschland junge Menschen mit Zuwanderungsgeschichte im Durchschnitt statistisch signifikant geringere Leseleistungen sowie Leistungen in den Naturwissenschaften aufweisen als junge Menschen ohne Zuwanderungsgeschichte (Autorengruppe Bildungsberichterstattung, 2022; Rauch, Mang, Härtig & Haag, 2016). Um diese Benachteiligungen auszugleichen, gerieten auf Ebene der Bildungspolitik zwei zentrale Forderungen in den Blick, die perspektivisch auf eine Verbesserung der schulischen Leistungen von Schüler:innen aus eingewanderten Familien abzielen: (1) eine verstärkte, fächerübergreifende Förderung bildungssprachlicher Kompetenzen von Schüler:innen bei gleichzeitiger Wertschätzung und Förderung der lebensweltlichen Mehrsprachigkeit der Schüler:innen (Kultusministerkonferenz, 2013, S. 7–8, 2019, S. 4–5) und (2) eine Erhöhung des Anteils an Lehrkräften mit Zuwanderungsgeschichte im deutschen Schulsystem (vgl. Rotter, 2012: 205; Rotter 2013: 32). Letztere stehen im deutschsprachigen Raum seit etwa zwei Jahrzehnten im bildungspolitischen Fokus (vgl. Ova, Stein & Zimmer, 2020; Rotter, 2012). Deren Anteil wurde an allgemeinbildenden Schulen im Jahr 2021 auf 13,3 Prozent geschätzt (vgl. Mediendienst Integration, 2022) und fällt damit deutlich hinter den Anteil an Schüler:innen mit Zuwanderungsgeschichte zurück. Angenommen wird, dass diese Lehrkräfte aufgrund geteilter biografischer Erfahrungen als Rollenvorbilder für Schüler:innen mit Zuwanderungsgeschichte dienen und damit zu einer verbesserten Integration und Leistungsfähigkeit derselben beitragen können (vgl. Rotter, 2012: 205; Ova et al. 2020; 113). Im Fokus der Bildungspolitik stand zudem von Beginn an die Erwartung, dass durch das “spezifische Kompetenzprofil” (Rotter, 2012: 205) von Lehrkräften mit Zuwanderungsgeschichte die Unterrichtsqualität nachhaltig verbessert werde (Presse- und Informationsamt der Bundesregierung, 2007, S. 117). In diesem Zusammenhang wird das Vorhandensein besonderer sprachlicher und sprachdidaktischer Kenntnisse sowie pädagogischer Fähigkeiten vorausgesetzt, die zu einer Kompensation der ausgemachten Benachteiligungen von Schüler:innen mit Zuwanderungsgeschichte beitragen sollen (Rotter 2012: 206). Bisher liegen für den deutschsprachigen Forschungskontext kaum Erkenntnisse darüber vor, inwieweit die bildungspolitisch formulierte Annahme einer stärkeren unterrichtlichen Berücksichtigung der sprachlichen Diversität der Schüler:innenschaft von Lehrkräften

mit Zuwanderungsgeschichte überhaupt zutrifft. Dies gilt vor allem für Schulen in sozialräumlich benachteiligten Lagen, an denen sich die kulturell-sprachliche Vielfalt der Schüler:innen in besonderem Maße niederschlägt.

Fragestellung

Der Beitrag untersucht die folgende Forschungsfrage: Wie gestaltet sich die sprachsensible Unterrichtspraxis von Lehrkräften mit und ohne Zuwanderungsgeschichte an Schulen in sozialräumlich benachteiligten Lagen?

Methode

Als Datenbasis dient eine standardisierte Online-Befragung von 911 Lehrkräften und pädagogischen Fachkräften aus 35 Schulen in benachteiligten Lagen in Nordrhein-Westfalen. Die Befragung wurde im Rahmen eines bundeslandweiten Modellprojekts zur Förderung von Schulen in sozialräumlich benachteiligten Lagen durchgeführt und erfasste die Ausgangssituation an den beteiligten Schulen, u. a. im Hinblick auf eine sprachbewusste Unterrichtspraxis der Lehrpersonen. Zur statistischen Auswertung wurden neben Gruppenvergleichen von Lehrkräften mit und ohne Zuwanderungsgeschichte Mehrebenenanalysen durchgeführt, in denen neben dem sprachlichen Hintergrund und Zuwanderungshintergrund der Befragten professionsbezogene und schulische Variablen berücksichtigt werden.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse verweisen auf statistisch signifikante Unterschiede zwischen Lehrkräften mit und ohne Zuwanderungsgeschichte mit Blick auf ihre sprachbewusste Unterrichtsgestaltung: Lehrkräfte mit Zuwanderungsgeschichte bezogen häufiger die Mehrsprachigkeit der Schüler:innen in den Unterricht ein und reicherten ihren Unterricht häufiger sprachlich an. In den multiplen Regressionsanalysen zeigte sich, dass diese Zusammenhänge teilweise durch schulische und individuelle Merkmale erklärt werden können. Unser Beitrag verweist auf die besondere Rolle von Lehrkräften mit Zuwanderungsgeschichte für die Förderung eines kritisch sprachbewussten Unterrichts an Schulen in benachteiligten Lagen und ordnet diese in die aktuelle bildungspolitische Diskussion ein.

Konnektoren in argumentativen Texten der 5. Klasse – welche, wie, wozu?

Elisabeth Mundt, Madeleine Domenech
Universität Kassel, Germany

Forschungsstand und Fragestellung

Der enge Zusammenhang von sprachlichen Kompetenzen und fachlichem Lernen ist gut belegt (z.B. Kempert at al., 2016; 2019). Dabei hat sich das sprachliche Register der sog. Bildungssprache als besonders bedeutsam erwiesen (Berendes at al., 2013; Haag et al., 2013; Heppt et al., 2016). Ein zentrales Element der Bildungssprache stellen Konnektoren dar – Konjunktionen, Subjunktionen oder Adverbien wie *sondern, sowie* oder *somit*, die Sätze, Satzteile und Wortgruppen auf bestimmte Art und Weise miteinander verbinden. Dass sich das Verständnis solcher Konnektoren auf das schulische Lernen auswirkt, konnte bereits für die Bereiche Lesen und Mathematik nachgewiesen

werden (Kohnen & Retelsdorf, 2019; Volodina et al., 2021). Auch für das Argumentieren ist die Kenntnis verschiedener Konnektoren notwendige Voraussetzung (Quasthoff & Domenech, 2016).

Um das Verständnis von Konnektoren gezielt trainieren zu können, bedarf es einer genauen Analyse des Status quo. Erste korpusanalytische Forschungen geben Einblick in den Konnektorengebrauch von Schüler*innen in einzelnen Domänen und Textsorten (z.B. Ricart Brede, 2014). Zur Verwendung von Konnektoren in argumentativen Zusammenhängen zu Beginn der Sekundarstufe 1 ist allerdings noch wenig bekannt. Auch zu den Zusammenhängen zwischen Konnektoren und spezifischen Argumentationsstrategien wie z.B. dem Einbezug und der Entkräftung von Gegenargumenten (z.B. Feilke, 2010; Rezat, 2011) mangelt es bisher an empirischer Evidenz. Diese Puzzleteile will die vorliegende Studie ergänzen: Welche Konnektoren werden in argumentativen Texten in Klasse 5 verwendet? Und wie hängt der Einsatz von Konnektoren mit dem Einbezug und der Entkräftung von Gegenargumenten zusammen?

Methode

Diese Fragen sollen mit Hilfe von Daten des BMBF-Projekts FUnDuS (Wild et al., 2018) beantwortet werden. Grundlage der Analysen bilden die Texte von 1448 Schüler*innen der 5. Klasse (571 Hauptschule, 877 Gymnasium), die eine Schreibaufgabe bearbeiteten, in der sie einen argumentativ-persuasiven Brief verfassen sollten. Neben der Textlänge wurden die argumentativen Konnektoren in diesen Texten mit guter Übereinstimmung geratet; dabei wurden Anzahl, Diversität (Anzahl verschiedener Konnektoren) und – bisher nur für eine Teilstichprobe (n = 200) – semantische Korrektheit der gebrauchten Konnektoren erfasst. Zudem wurde erhoben, ob die Schreiber*innen die Strategie des Einbezugs und der Entkräftigung von Gegenargumenten (vgl. TexQu-Ap, Quasthoff & Domenech, 2016) verwendeten (0 nein, 1 ja).

Ergebnisse

Im Durchschnitt schrieben die Schüler*innen Texte mit einer Länge von 59.3 Wörtern (SD = 26.8, Min = 2, Max = 177) und verwendeten darin 1.2 argumentative Konnektoren (SD = 1.2, Min = 0, Max = 9), wobei die mittlere Diversität M = 1.1 (SD = 0.9, Min = 0, Max = 5) betrug. Insgesamt wurden 20 verschiedene kausale, adversative, konzessive und finale Konnektoren verwendet. Am häufigsten waren dabei die Wörter *weil* (638) und *aber* (407), am seltensten die Wörter *dagegen* (2), *allerdings* (1) und *dennoch* (1).

Die Anzahl sowie die Diversität der verwendeten Konnektoren korrelieren erwartungskonform mit der Textlänge (r_A = .45, p < .001; r_D = .41, p < .001) und weisen auch mit der Argumentationsstrategie der Berücksichtigung von Gegenargumenten Zusammenhänge auf (r_A = .26, p < .001; r_D = .28, p < .001).

Erste Analysen zur semantischen Korrektheit der eingesetzten Konnektoren deuten darauf hin, dass bei einzelnen Konnektoren noch Unsicherheiten bzgl. ihrer Bedeutung bestehen; nach einem ersten Rating einer Teilstichprobe liegt die Fehlerquote bei dem hoch frequenten adversativen Konnektor *aber* mit 20% besonders hoch.

Diskussion und Ausblick

Die Ergebnisse geben einen guten Einblick in die aktive Nutzung von Konnektoren bei argumentativ-persuasiven Schreibaufgaben zu Beginn der Sekundarstufe 1. Dabei gibt es erste Hinweise darauf, dass ein explizites Training mancher Konnektoren sinnvoll sein könnte. Der erwartete Zusammenhang zwischen dem Konnektorengebrauch und der Berücksichtigung von Gegenargumenten lässt sich finden, auch wenn er – wahrscheinlich auch aufgrund der grobkörnigen Erfassung – relativ gering ausfällt.

In weiteren Analysen soll die Ermittlung der semantischen Fehlerquote für die gesamte Stichprobe finalisiert und abgesichert werden. Außerdem soll überprüft werden, ob die Konnektoren bestimmter semantischer Kategorien (z.B. adversativ oder konzessiv) besonders eng mit Argumentationsstrategien wie dem Einbezug von Gegenargumenten zusammenhängen.

Ist sprachsensibler Unterricht nachweislich wirkungsvoll? Ergebnisse eines systematischen Reviews

Leonie Regina Twente[1,2], Tetyana Vasylyeva[1], Till Woerfel[1]
[1]Universität zu Köln, Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache; [2]Universität zu Köln, Philosophische Fakultät, Institut für deutsche Sprache und Literatur II

Unter dem Begriff „Sprachsensibler Unterricht“ werden verschiedene didaktisch-methodische Ansätze zusammengefasst, die explizit Fach- und Sprachlernen verknüpfen und dadurch eine Verbesserung fachlicher und (fach-)sprachlicher Fertigkeiten aller Lernenden im Kontext sprachlich heterogener Lerngruppen erzielen sollen (vgl. Becker-Mrotzek & Woerfel, 2020). In Abgrenzung zu anderen sprach-und-fachintegrierten Unterrichtsansätzen (z. B. CLIL) sind beim sprachsensiblen Unterricht Ziel- und Amts- bzw. Unterrichtssprache identisch.

Sprachsensibles Unterrichten ist in Deutschland curricular verankert (Becker-Mrotzek et al., 2021, S. 250) und wurde in den letzten Jahren vermehrt als Gegenstand von empirischen Studien aufgegriffen. Allerdings fehlte bislang eine systematische Zusammenschau der Evidenz zur Wirkung sprachsensiblen Unterrichts, die die gesamte Bandbreite existierender Ansätze berücksichtigt (Woerfel et al., 2020, S. 9). Vor dem Hintergrund der Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring (Sekretariat der Kultusministerkonferenz, 2016) ist es Aufgabe der Wissenschaft, belastbare Ergebnisse zur Wirkung sprachsensiblen Unterrichts vorzulegen. Eine dafür geeignete Methode ist das systematische Review, dessen Ziel die Identifikation, Bewertung und Synthese der verfügbaren Evidenz zu einer bestimmten Fragestellung ist (Woerfel et al., 2021, S. 163–165).

Mit dem Abschluss des Projekts „Systematisches Review zur Wirkung von didaktisch-methodischen Ansätzen des sprachsensiblen Unterrichts“ liegt nun eine Zusammenschau der aktuellen internationalen Forschungslage vor, die die folgende Forschungsfrage beantwortet:

Sind sprachsensible Unterrichtsansätze im Primar- und Sekundarbereich nachweislich wirkungsvoller als nicht-sprachsensibler Fachunterricht?

Das Vorgehen wurde entlang der aktuellen Standards für systematische Reviews entwickelt (u. a. *EPPI-Centre, PRISMA*). Aufgrund der zu erwarteten Heterogenität und Komplexität der Primärstudien wurde eine Synthese ohne Metaanalyse unter Berücksichtigung aktueller Guidelines (u. a. Campbell et al., 2020; Petticrew & Roberts, 2006) durchgeführt.

Die in einem dreistufigen Verfahren mit Text Mining (Twente & Woerfel, 2020) ermittelten Suchbegriffe wurden zu Suchanweisungen für sechs elektronische Datenbanken kombiniert. Zusätzlich wurden Handsuchen, ein Call for Preprints und eine Expert:innenabfrage durchgeführt. Im Zeitraum von 1990-2021 wurden 3016 englisch- und deutschsprachige Dokumente ermittelt und entlang der vorab definierten Inklusionskriterien gescreent. Die identifizierten Studien (n=55) wurden mittels dafür entwickelter Kriterienraster hinsichtlich ihrer internen und externen Validität bewertet; zwei Studien wurden wegen Nichterfüllung der Kriterien ausgeschlossen. Die in die Synthese eingeschlossenen Studien (n = 53) umfassen experimentelle, quasi-experimentelle und nicht-experimentelle Studiendesigns mit inferenzstatistischer Auswertung und wurden in neun verschiedenen Ländern durchgeführt.

Die 53 Studien überprüfen 30 verschiedene sprachsensible Ansätze, die sich folgenden Sprachschwerpunkten zuordnen lassen: Lesen, Schreiben, Mündlichkeit, Wortschatz, übergreifende Ansätze (mehrere Sprachschwerpunkte). Ansätze zu mehreren Schwerpunkten sowie zu Lesen wurden häufiger in mehreren Studien evaluiert, während Ansätze in den Bereichen Mündlichkeit, Wortschatz und Schreiben nur in je einer einzigen Studie evaluiert wurden.

Der Vortrag präsentiert die zentralen Ergebnisse des Reviews: Die Synthese der inkludierten Primärstudien ergibt, dass Schülerinnen und Schüler des Primar- und Sekundarbereichs, die sprachsensibel unterrichtet werden, besser, mindestens aber vergleichbar sprachliche und/oder fachliche Lernziele erreichen als solche, die nicht sprachsensibel unterrichtet werden. Die Studien sind jedoch äußerst heterogen, u.a. hinsichtlich der Populationen, Outcomes, Kontrolltreatments und Interventionen.

Im Vortrag diskutieren wir, inwiefern die Methode des systematischen Reviews dazu beiträgt, vorläufiges Wissen in einigermaßen gesichertes Wissen für Forschung und Praxis zu überführen. Dabei nehmen wir in den Blick, wie wichtig es ist, die unterschiedliche Qualität und Generalisierbarkeit der Primärstudien bei der Interpretation der Ergebnisse zu berücksichtigen. Außerdem sprechen wir Implikationen für die Forschung, die Bildungspraxis sowie die Zusammenarbeit verschiedener Stakeholder an.

Textentlastung mithilfe empirisch fundierter Regeln Leichter Sprache

Sarah Bormann, **Katrin Böhme**
Universität Potsdam, Germany

Schulisches Lernen ist in der Regel sprachlich basiert. So werden bspw. für die Wissensvermittlung im Fachunterricht nach wie vor Sachtexte als zentrales Medium eingesetzt (Ahrenholz et al., 2019). Oftmals überfordert jedoch die Komplexität von Schulbuchtexten Schüler:innen mit gering ausgeprägter Lesekompetenz (u.a. Benholz & Rau, 2011; Oleschko & Moraitis, 2012). Folglich führt eine gering ausgeprägte Lesekompetenz potentiell zu Lernschwierigkeiten in allen Fächern. Eine Möglichkeit, das Leseverstehen schwacher Leser:innen zu unterstützen und damit inhaltliches Lernen auch dieser Zielgruppe zu ermöglichen, ist die Reduzierung der Textkomplexität auf der sprachlichen Ebene (Schneider et al., 2018). Bislang ist allerdings noch nicht ausreichend empirisch geklärt, welche konkreten Prinzipien tatsächlich zuverlässig zu einer verständnisfördernden sprachlichen

Entlastung führen (Schneider et al., 2018). Entsprechende Untersuchungen nutzen jeweils verschiedene Entlastungsprinzipien und liefern ein heterogenes Befundbild (Härtig et al., 2019; Schneider et al., 2018; Apolin, 2002).

Seit einigen Jahren wird die regelbasierte schriftsprachliche Varietät Leichte Sprache als ein mögliches Entlastungsprinzip diskutiert, das im schulischen Kontext zur Unterstützung von schwachen Leser:innen beitragen kann (Bormann & Böhme, 2021). Diesbezüglich gibt es jedoch einen kontroversen fachdidaktischen Diskurs: Werden Texte, die nach Regeln Leichter Sprache verfasst wurden, von leseschwächeren Schüler:innen tatsächlich besser verstanden als standardsprachliche Texte (Bormann & Böhme, 2021)? Ist Leichte Sprache überhaupt für den Schulkontext geeignet oder bedeutet sie eine zu starke Vereinfachung (Lasch, 2017)? Welche potentiell negativen Effekte, wie bspw. eine Stigmatisierung von Schüler:innen oder eine Abkehr von zentralen Bildungszielen kann der Einsatz Leichter Sprache im schulischen Kontext mit sich bringen (Kilian, 2017; Lasch, 2017)?

Der vorliegende Beitrag führt die genannten Problemstellungen zusammen und geht der Frage nach, ob empirisch fundierte Regeln Leichter Sprache tatsächlich für eine verständnisfördernde Textentlastung genutzt werden können und welchen Einfluss die Jahrgangsstufe sowie die Lesekompetenz der Schüler:innen auf das Textverständnis entlasteter Texte haben. Im Projekt „Lies entlastet – Textverständnis entlasteter Schulbuchtexte“ (LiesE) wurde die Textverständlichkeit von standardsprachlichen Schulbuchtexten der Fächer Sachunterricht, Naturwissenschaften und Gesellschaftswissenschaften mit der Textverständlichkeit von entlasteten Versionen dieser Schulbuchtexte verglichen. Darüber hinaus wurde untersucht, ob der entlastende Effekt der Textversion je nach Jahrgangsstufe und Lesekompetenz der Schüler:innen unterschiedlich ausfällt. Hierfür lasen 198 Schüler:innen, davon 50% Mädchen, der dritten bis sechsten Jahrgangsstufe in einem within-subject-Design je zwei standardsprachliche und zwei entlastete Texte zu verschiedenen Themen und absolvierten nach dem Lesen jedes Textes einen schriftlichen Verständnistest. Das Durchschnittsalter der Schüler:innen betrug 10,7 Jahre, der Anteil von Schüler:innen mit einer anderen Familiensprache als Deutsch war mit ca. 3% sehr gering. Zusätzlich zum Textverständnis wurden basissoziodemografische Hintergrundinformationen und Einstellungen zum Lesen in einem kurzen Schülerfragebogen sowie die allgemeine Lesekompetenz mithilfe des standardisierten Lesetests ELFE II (Lenhard, Lenhard & Schneider, 2017) erhoben. Die Daten wurden mithilfe einer mixed Anova mit Messwiederholung ausgewertet. Es zeigte sich, dass die entlasteten Texte für alle Schüler:innen besser verständlich waren als die standardsprachlichen Schulbuchtexte, also zu signifikant besseren Leistungen im schriftlichen Verständnistest führten. Statistisch signifikante Interaktionseffekte zwischen der Textversion und der Jahrgangsstufe sowie der Lesekompetenz zeigten sich hingegen nicht. Die Ergebnisse der Studie deuten darauf hin, dass solche Regeln Leichter Sprache, deren verständnisfördernde Wirkung bereits empirisch belegt ist, tatsächlich gewinnbringend zur Entlastung von Sachtexten für den Schulunterricht genutzt werden können. Hierbei waren sprachlich weniger komplexe Versionen der Schulbuchtexte für die Schüler:innen aller Jahrgangsstufen und Kompetenzniveaus besser verständlich als die Originaltexte aus den Schulbüchern. Kompensatorische Effekte spezifisch für schwache Leser:innen zeigten sich nicht. In

Anknüpfung an diese Befunde werden im Beitrag Implikationen und Bedingungen für die schulische Umsetzung von Textentlastung auf der Basis von Regeln Leichter Sprache diskutiert.

B-08: Vor- und Grundschule

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S05 T00 B71
Chair der Sitzung: Nadine Poltz

Wer profitiert nicht von einer Förderung der mathematischen Kompetenzen im Kindergarten?

Nadine Poltz, Luisa Wagner
University of Potsdam, Germany

Bereits vor Schuleintritt zeigen Kinder interindividuelle Unterschiede in ihren numerischen Fertigkeiten (z.B. Aunola et al., 2004). Diese Fertigkeiten gelten als die zentralen Vorläuferfertigkeiten späterer schulischer Rechenleistungen (Nguyen et al., 2016). Darüber hinaus konnte gezeigt werden, dass auch die sprachlichen Fähigkeiten auf die Kompetenzentwicklung im Rechnen wirken (Prediger, 2019). Daneben erwiesen sich auch die Aufmerksamkeitsleistungen (z.B. Knievel et al., 2010) sowie die visuell-räumlichen Wahrnehmungsfertigkeiten (z.B. Mazzocco & Myers, 2003) als relevant für den Erwerb von Rechenfertigkeiten.

In verschiedenen Studien konnte gezeigt werden, dass die Förderung der Fertigkeiten im vorschulischen Bereich eine positive Wirkung auf die Lernentwicklung in der Grundschule zeigt und somit das Risiko für Lernschwierigkeiten verringert werden kann (z.B. Moraske et al., 2018).

Allerdings zeigen sich in Studien immer wieder Kinder, die auf die jeweiligen Fördermaßnahmen nicht ansprechen. Diese Kinder werden als Non-Responder bezeichnet. In den letzten Jahren rückten sie in den Fokus, um herauszufinden, warum sie nicht von der Förderung profitieren und welche Merkmale einen Einfluss darauf ausüben, um Fördermaßnahmen anpassen zu können.

Al Otaiba und Fuchs (2002) identifizierten acht Studien, die nach Ursachen für die Nicht-Ansprache auf Fördermaßnahmen suchten. Es kristallisierten sich phonologische Bewusstheit, Intelligenz, Arbeitsgedächtnis, RAN sowie Verhaltensauffälligkeiten und sozioökonomischer Status als Einflussfaktoren heraus.

Auffallend ist, dass die Untersuchungen fast ausschließlich für den schriftsprachlichen Bereich durchgeführt wurden und Betrachtungen der Non-Responder im mathematischen Kontext fehlen. Daher soll sich dieser Beitrag diesen Non-Respondern widmen:

1a: Inwiefern unterscheiden sich Non-Responder von Respondern in den Leistungs- und Verhaltensvariablen zu t1?

1b: Welche Variable zu t1 sagt die Responsiveness hauptsächlich vorher?

2: Wie entwickeln sich Non-Responder im Vergleich zu Respondern in den ersten Schuljahren?

Methodik: In der SCHUES-Studie wurden 40 Vorschulkinder (*M*=61,65 Monate) mit einem Risiko einer Rechenstörung über elf Wochen im Zahlen- und Mengenverständnis gefördert (s. Moraske et al., 2018). Zu t1 wurden das Zahlen- und Mengenverständnis, die verbale und nonverbale Intelligenz, die selektive Aufmerksamkeit, die expressive und rezeptive Sprache sowie die phonologische Bewusstheit erhoben. Daneben wurden die Leistungen in den Bereichen der visuell-räumlichen Wahrnehmung, des Arbeitsgedächtnisses und des RAN erhoben. Zusätzlich wurden Verhaltenseinschätzungen der Eltern zu externalisierenden Verhaltensauffälligkeiten und Aufmerksamkeitsproblemen erfasst. Alle Kinder wurden einmal jährlich über vier Jahre in ihren mathematischen Kompetenzen untersucht.

Zu t2 wurden 15 Non-Responder und 25 Responder identifiziert. Die Responder unterschieden sich von den Non-Respondern weder nach Geschlecht ($\chi^2(1)=0{,}07$, $p=.870$), höchstem beruflichen Qualifikationsniveau der Eltern ($\chi^2(4)=7{,}32$, $p=.120$), der gesprochen Sprache zu Hause ($\chi^2(2)=1{,}89$, $p=.390$), in der Anzahl teilgenommener Sitzungen ($t(35)=0{,}45$, $p=.653$) und des Alters ($t(38)=-1{,}59$, $p=.121$).

Ergebnisse: Responder unterschieden sich von Non-Respondern signifikant (nach Bonferroni-Korrektur auf $p<.004$) in den folgenden Variablen: verbale Intelligenz, expressive Sprache und phonologische Bewusstheit. Tendenziell zeigten Non-Responder auch in den Variablen nonverbale Intelligenz, rezeptive und expressive Sprache, visuell-räumlicher Notizblock und zentrale Exekutive geringe Leistungen sowie ein höheres Ausmaß an externalisierenden Verhaltensauffälligkeiten ($p<.05$). In einer schrittweisen Diskriminanzanalyse ($\lambda=.657$, $\chi^2(1)=15{,}75$, $p<.001$) zeigte die verbale Intelligenz als einzige Variable einen signifikanten Beitrag. Während Responder in den Klassen 1-3 im Mittel durchschnittliche Leistungen im mathematischen Bereich erzielten, fielen die Leistungen der Non-Responder in den auffälligen Bereich. Statistisch bedeutsam wurde dieser Gruppenunterschied nur für die erste Klasse ($t(17)=3{,}58$, $p<.001$).

Diskussion:

Die Ergebnisse zeigen, dass die sprachliche Intelligenz entscheidend dazu beiträgt, ob eine mathematische Intervention erfolgreich ist. Die sprachlichen Faktoren wirken folglich nicht nur auf Förderung im schriftsprachlichen Bereich, sondern auch auf die Responsiveness in der mathematischen Förderung. Daneben zeigt sich, dass die Non-Responder bis in die Grundschulzeit geringere mathematische Kompetenzen aufweisen, obwohl sie eine Förderung durchliefen. Dies zeigt sehr deutlich, dass Kinder mit zusätzlichen sprachlichen Auffälligkeiten, insbesondere der sprachlichen Intelligenz als kognitive Basisfunktion, von mathematischen Förderungen allein nicht profitieren. Für diese Kinder erscheinen ganzheitliche Förderstrategien angeraten, die neben den mathematischen Kompetenzen auch sprachliche Fertigkeiten fokussieren.

Validierung eines Erfassungsinstruments für Strategiewissen zu selbstreguliertem Lernen bei Vorschulkindern

Lena Grüneisen, Laura Dörrenbächer-Ulrich, Franziska Perels
Universität des Saarlandes, Germany

Das selbstregulierte Lernen (SRL) als Fähigkeit, die eigenen Gedanken und Verhaltensweisen anhand individueller Lernziele selbstgerichtet zu steuern (Zimmerman, 2011), ist bereits im Vorschulalter ein signifikanter querschnittlicher wie auch längsschnittlicher Prädiktor für akademische Kompetenz (Gestsdottir et al., 2014; Willoughby et al., 2011). Entsprechend ist es wichtig, diejenigen Vorschulkinder zu identifizieren, die Defizite im SRL aufweisen, um ihnen mithilfe einer gezielten SRL-Förderung den Schuleinstieg erleichtern zu können. Häufig werden zur vorschulischen SRL-Diagnostik Fremdurteile durch Bezugspersonen oder Beobachtungsverfahren genutzt (z.B Dörrenbächer & Perels, 2018; McDermott et al., 2002; Whitebread et al., 2009). Diese Messverfahren erlauben jedoch keine Aussagen zu ausschließlich internalen und daher nicht beobachtbaren SRL-bezogenen Prozessen, wie zum Beispiel dem Wissen über SRL-Strategien.

Im Rahmen der derzeit laufenden Studie soll daher ein Test zur direkten Erfassung des Wissens über SRL-Strategien bei Vorschulkindern validiert werden. Der Test stellt eine Überarbeitung des von Jacob et al. (2019) entwickelten und validierten Messverfahrens dar. Den theoretischen Rahmen bildet eine für das Vorschulalter adaptierte Version des sozial-kognitiven Modells von Zimmerman (2000), das SRL als zyklischen, in Handlungsphasen unterteilten Prozess betrachtet. Das Verfahren ist als den Vorschulkindern vorgelesene und mit Bildern illustrierte Geschichte konzipiert, in der der Protagonist *vor*, *während* und *nach* einer zielorientierten Handlung mit Situationen konfrontiert ist, in denen das Anwenden von SRL-Strategien hilfreich ist. Zu jeder Situation werden als Items zwei Handlungsalternativen vorgelesen; eine davon stellt die Nutzung einer SRL-Strategie dar, die andere non-selbstregulatives Verhalten. Die Kinder sollen durch Zeigen auf eine visuelle Antwortskala die Zieldienlichkeit der Handlungsalternativen beurteilen. Ihre Leistung ergibt sich aus der Übereinstimmung ihres Urteils mit der tatsächlichen SRL-Relevanz des jeweiligen Items.

Im Zuge der Weiterentwicklung des zu revidierenden Verfahrens von Jacob et al. (2019) erfolgte eine Überarbeitung des Protagonisten der dem Test immanenten Geschichte sowie eine Anpassung der den Kindern präsentierten Antwortskala. Des Weiteren wurden einzelne Items umfomuliert, herausgenommen oder neu generiert. Diese Überarbeitungsmaßnahmen sollen der Verringerung einer Ja-sage-Tendenz, der Erhöhung der Reliabilität und hierausfolgend der Validität des Messinstruments dienen. Der überarbeitete Strategiewissenstest wird derzeit unter Einsatz einer umfangreichen Testbatterie in einer geplanten Stichprobe von $N = 100$ Vorschulkindern validiert. Die Testbatterie beinhaltet zur Berechnung der konvergenten Validität zwei direkte und ein indirektes Messverfahren für selbstregulative Kompetenzen (Geschenkverpackungsaufgabe zur Erfassung „heißer“ Selbstregulation, Kochanska et al., 2000; Test zum metakognitiven Wissen zur Erfassung metakognitiver SRL-Strategien, Lockl et al., 2016; COMPSCALE als Fremdurteil zum SRL-Verhalten, Lange et al., 1989). Als divergentes Maß wird die Psychomotorik erhoben (Adaption des Subtests Grobmotorik der IDS-2, Grob & Hagmann-von Arx, 2018; eine an den Test Psychomotorik der IDS-2 angelehnte, selbstgenerierte Fremdurteilskala zur motorischen Entwicklung). Zur Berechnung der Kriteriumsvalidität werden akademische Kompetenzen erfasst (Test Logisch-Mathematisches Denken der IDS-2, Grob & Hagmann-von Arx, 2018; selbstgenerierte Fremdurteilskala zum erwarteten Schulerfolg). Als potenzielle Kontrollvariable wird das schlussfolgernde Denken als Intelligenzmaß erhoben (Adaption der Coloured Progressive Matrices, Raven, 1998). Für die Datenanalyse sind auf Item-Ebene Analysen der Item-Schwierigkeiten und -Trennschärfen geplant. Auf Ebene des gesamten Messverfahrens soll eine Reliabilitäts- und Validitätsanalyse erfolgen. Als Ergebnisse der Validitätsanalyse werden hohe und signifikante Korrelationen zwischen dem revidierten Strategiewissenstest und den konvergenten sowie den Kriteriumsmaßen erwartet. Zu den divergenten Maßen werden hingegen niedrigere und nicht-signifikante Korrelationen erwartet.

Die Ergebnisse einer vorläufigen Datenanalyse zur bisherigen Stichprobe ($n = 51$ Vorschulkinder) sprechen bis auf eine Ausnahme für die Annahme der Hypothesen. Lediglich zum Test Logisch-Mathematisches Denken konnte entgegen der Hypothesen keine signifikante Korrelation gefunden werden. Dies ist möglicherweise dadurch bedingt, dass in der Vorschule noch kein Mathematikunterricht stattfindet und die vorschulischen mathematischen Kompetenzen daher

womöglich noch unbeeinflusst vom kindlichen (zielgerichteten) Lernverhalten sind. Bis zur zehnten GEBF-Tagung sollten die finalen Ergebnisse vorliegen, die dann vorgestellt und diskutiert werden können.

Lernen aus Fehlern im Vor- und Grundschulalter

Carolin Burmeister, Robert Grassinger
Pädagogische Hochschule Weingarten, Deutschland

Theoretischer Hintergrund

Lernprozesse von Kindern werden vielfach durch das Anstoßen eines ‚trial & error' angeregt und sind besonders dann einem Wissensaufbau zuträglich, wenn Kinder aus Fehlern lernen (Hascher & Hagenauer, 2010; Spychiger et al., 2006). Vor- und Grundschulkinder erhalten durch Erzieher*innen und Lehrkräfte Fehlerrückmeldungen, die besonders dann als fruchtbar für eine Veränderung von Wissensstrukturen erscheinen, wenn Lernende affektiv-motivational adaptiv und handlungsadaptiv auf Fehlerfeedback reagieren (Dresel et al., 2013; Steuer et al., 2013). Fehlerreaktionen werden beeinflusst von verschiedenen individuellen Determinanten wie dem Fähigkeitsselbstkonzept oder selbstregulativen Kompetenzen (ebd.). Empirische Evidenz für diese Annahmen liegt für Vor- und Grundschulkinder kaum vor, jedoch für Sekundarstufenschüler*innen (Dresel et al., 2013; Grassinger & Dresel, 2017; Grassinger et al., 2015; Steuer, 2014; Tulis et al., 2016).

Angesichts sich stark entwickelnder Kompetenzen und entwicklungspsychologischer Veränderungen individueller Determinanten für ein Lernen aus Fehlern (Artelt et al., 2012; Fischer & Rustemeyer, 2007; Holodynski et al., 2013; Lockl & Schneider, 2003; Spinath & Spinath, 2005), lassen sich empirische Befunde von Sekundarstufenschüler*innen nicht per se auf Vor- und Grundschulkinder übertragen. Nach Tulis und Kolleg*innen (2016) sind auch kontextuelle Determinanten wie das Fehlerklima hinsichtlich Fehlerreaktionen postuliert. Empirische Evidenz für die Wahrnehmung des Fehlerklimas im Klassenzimmer liegt ebenfalls vorwiegend für den Sekundarstufenbereich vor (Steuer, 2014; Steuer et al., 2013) und kann aufgrund der unterschiedlichen Gestaltung von Lernprozessen nicht einfach von der Sekundarstufe auf Kindertageseinrichtungen und Grundschulen abgeleitet werden.

Ziel der Arbeit ist, ein Lernen aus Fehlern von Vor- und Grundschulkindern unter Berücksichtigung individueller und kontextueller Bedingungen für ein Lernen aus Fehlern zu untersuchen.

Fragestellung

Im Fokus der Arbeit steht die Frage, welche individuellen und kontextuellen Determinanten ein Lernen aus Fehlern von Vor- und Grundschulkindern beeinflussen.

Methode

Es wurden 600 Grundschüler*innen der ersten, zweiten und dritten Jahrgangsstufe anhand eines Fragebogens hinsichtlich ihrer Tendenzen, adaptiv auf Fehler zu reagieren, befragt. Der Fragebogen beinhaltete überdies Items zum Interesse am Fach Mathematik, zum Fähigkeitsselbstkonzept in Mathematik, zur Einstellung zu Fehlern und zur Wahrnehmung des Fehlerklimas im Klassenzimmer sowie Subtests zur Erfassung kognitiver Fähigkeiten und metakognitiven Wissens der Schüler*innen. Zudem wurden adaptive Fehlerreaktionen, das Fähigkeitsselbstkonzept, kognitive Fähigkeiten,

metakognitives Wissen und die Wahrnehmung des Fehlerklimas im Vorschulprogramm von 41 Vorschulkindern anhand eines spielbasierten standardisierten Interviews erfasst. Berechnet wurden überwiegend lineare und multiple Regressionen.

Ergebnisse

Die Ergebnisse konstatieren, dass Vorschulkinder nach einem Fehler überwiegend die Lernfreude aufrechterhalten und adaptiv auf Fehler reagieren, wobei selbstregulativen Kompetenzen eine bedeutsame Rolle zukommt. Insbesondere kognitive Fähigkeiten und das Fehlerklima im Vorschulprogramm beeinflussen ein Lernen aus Fehlern im Vorschulalter.

Mit dem Eintritt in die Primarstufe zeigen Schüler*innen affektiv-motivational adaptive und handlungsadaptive Fehlerreaktionen, die sich bis in die dritte Klasse – entgegen entwicklungspsychologischer Veränderungen und der Ausdifferenzierung von Kompetenzen – nicht modifizieren. Das heißt, es lassen sich bei Lernenden unabhängig von der Klassenstufe ähnliche adaptive Fehlerreaktionen feststellen. Wesentlich beeinflusst wird ein Lernen aus Fehlern ab der Grundschule bis in die dritte Klasse durch das Interesse am Fach Mathematik, die Einstellung zu Fehlern, das Fähigkeitsselbstkonzept in Mathematik und das Fehlerklima im Klassenzimmer.

Die Ergebnisse liefern empirische Evidenz dahingehend, dass Vor- und Grundschulkinder adaptive Fehlerreaktionen zeigen, die verstärkt durch das Interesse am Fach Mathematik, die Einstellung zu Fehlern und ein positives Fehlerklima im Klassenzimmer bzw. Vorschulprogramm begünstigt werden. Diese individuellen Determinanten und das positive Fehlerklima als Determinanten für ein Lernen aus Fehlern manifestieren sich als variable Faktoren, die auch durch die Lehrkraft steuerbar sind (Hartinger, 1997; Hascher & Hagenauer, 2011).

Mit Citizen Science Wissenschaft verstehen: Eine Intervention zur Förderung von wissenschaftlichem Denken bei Grundschulkindern

Julia Maria Lange[1], Julia Schiefer[1], Carola Garrecht[2], Markus Kleinhansl[1], Steffen Seitz[3], Thomas Scholten[3], Jessika Golle[1], Benjamin Nagengast[1], Ulrich Trautwein[1]

[1]Hector-Institut für Empirische Bildungsforschung, Universität Tübingen; [2]Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel; [3]Bodenkunde und Geomorphologie, Universität Tübingen

Theoretischer Hintergrund und Fragestellung

Wissenschaftliches Denken ist eine Schlüsselkompetenz des 21. Jahrhunderts (Fischer et al., 2014; Koerber & Osterhaus, 2021) und ein wesentliches Ziel des naturwissenschaftlichen Unterrichts (NRC, 2012; OECD, 2019). Zentrale Bereiche wissenschaftlichen Denkens sind ein adäquates Verständnis der Natur der Naturwissenschaften und der Methoden wissenschaftlicher Erkenntnisgewinnung (z.B. Kenntnis möglicher experimenteller Studiendesigns) sowie die Fähigkeit, Experimente durchzuführen, auszuwerten und wissenschaftliche Schlussfolgerungen zu ziehen (Zimmerman, 2007).

Im Rahmen eines landesweiten Enrichment-Programms für begabte Grundschulkinder in Baden-Württemberg (*Hector Kinderakademien*, vgl. Golle et al., 2018) wurde eine 8-wöchige Intervention zum Thema „Boden“ entwickelt. Ziel der Intervention ist, das wissenschaftliche Denken von Dritt- und Viertklässler:innen und ihre Motivation, sich mit naturwissenschaftlichen Themen auseinanderzusetzen, zu fördern.

Der Kurs enthielt Elemente des forschenden Lernens (Pedaste et al., 2015) wie die Planung eigener bodenkundlicher Experimente sowie die aktive Beteiligung an einem Citizen-Science-Projekt zum Thema „Boden“. Da sich Citizen-Science-Projekte mit offenen Forschungsfragen befassen, bieten sie Schüler:innen innovative, authentische und sinnstiftende Lerngelegenheiten (Bonney et al., 2016). Wir erwarteten, dass sich durch eine Teilnahme das wissenschaftliche Denken, die Motivation für Naturwissenschaften und das Inhaltswissen zu den Themen „Citizen Science“ und „Boden“ steigern lässt.

Methode

Stichprobe und Design

An unserer Intervention im Schuljahr 2021/22 nahmen *N*=98 Grundschulkinder (56.12% Jungen; Alter: *M*=9.01, *SD*=0.66; Klasse 2: *N*=2, Klasse 3: *N*=55, Klasse 4: *N*=41) teil. Die Intervention wurde von sechs Kursleitungen in Kleingruppen an sieben Akademien durchgeführt. Der vorliegenden präregistrierten Studie[1] liegt ein randomisiertes Warte-Kontrollgruppen-Design mit Prä- und Posttesterhebung zugrunde.

Instrumente

Wissenschaftliches Denken. Um die verschiedenen Bereiche wissenschaftlichen Denkens abzubilden, wurden die Experimentierfähigkeiten der Schüler:innen (Phan, 2007), ihr Verständnis der Natur der Naturwissenschaften (wie der Herkunft naturwissenschaftlichen Wissens, adaptiert nach Kremer, 2010) und ihr Verständnis experimenteller Studiendesigns (Osterhaus et al., 2015) erfasst.

Motivation. Um die Motivation der Schüler:innen zur Beschäftigung mit naturwissenschaftlichen Themen zu erheben, wurden ihr forschendes Interesse (adaptiert nach Holland, 1994), ihre Science Identity („Wissenschaftsidentität“, adaptiert nach Vincent-Ruz & Schunn, 2018) und ihr Selbstkonzept in Bezug auf die Durchführung von Experimenten (adaptiert nach Gaspard et al., 2021) erfasst.[2]

Inhaltswissen. Das Wissen der Schüler:innen zu den Themen „Boden“ und „Citizen Science“ wurde mit selbstentwickelten Items erhoben.

Statistische Analyse

Die Effekte der Intervention wurden mittels multipler Regressionen in R (R Core Team, 2022) analysiert (AVs: standardisierte Posttestwerte, UVs: Gruppenzuweisung [Interventionsgruppe=1, Kontrollgruppe=0], standardisierte Prätestwerte) und können aufgrund der Standardisierung der AVs und UVs als Effektstärken (Cohens d) interpretiert werden.

Ergebnisse

Wissenschaftliches Denken. Unsere vorläufigen Ergebnisse zeigen, dass Schüler:innen in der Interventionsgruppe im Vergleich zur Kontrollgruppe signifikant höhere Experimentierfähigkeiten (*B*=0.36, *p*=.006) entwickelten. Keine signifikanten Interventionseffekte konnten für das Verständnis der Herkunft naturwissenschaftlichen Wissens (*B*=0.29, *p*=.121) und das Verständnis experimenteller Studiendesigns beobachtet werden (*B*=0.1, *p*=.199).

Motivation. Unsere Ergebnisse deuten darauf hin, dass die Schüler:innen in der Interventionsgruppe im Vergleich zur Kontrollgruppe eine signifikant höhere Science Identity entwickelten (*B*=0.29,

p=.008). Kein signifikanter Interventionseffekt konnte für das forschende Interesse (*B*=-0.03, *p*=.43) und das Selbstkonzept in Bezug auf die Durchführung von Experimenten (*B*=-0.02, *p*=.454) nachgewiesen werden.

Inhaltswissen. Es zeigten sich signifikant positive Interventionseffekte für das Wissen über den Boden (*B*=1.16, *p*<.001) und Citizen Science (*B*=0.29, *p*<.001).

Die vorliegende Studie liefert damit Hinweise, dass die Intervention das wissenschaftliche Denken der Schüler:innen, ihre Science Identity sowie ihr Wissen über den Boden und Citizen Science fördert. Wissenschaftliches Denken ist eine wesentliche Voraussetzung für die sachgerechte Einordnung von Informationen und das Treffen wissenschaftlich fundierter Entscheidungen – Fähigkeiten, die in der Informationsgesellschaft von zentraler Bedeutung sind.

[1] Registry ID: 11380.1v1, Registry of Efficacy and Effectiveness Studies (https://sreereg.icpsr.umich.edu/sreereg/).

[2] In diesem Abstract konzentrieren wir uns auf eine Auswahl der Instrumente zur Messung von Motivation. Unsere Präregistrierung enthält eine detaillierte Beschreibung aller Instrumente.

B-09: Studienwahl und Studienmotivation

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S05 T00 B83
Chair der Sitzung: Thomas Niewöhner

Studienmotivation – Validierung und Normierung eines Fragebogens aus Perspektive der Erwartungs-Wert Theorie

Lucas Wloch[1], Theresa Schnettler[2], Sophie von der Mülbe[3], Lisa Bäulke[4], Markus Dresel[3], Carola Grunschel[1]
[1]Universität Münster, Germany; [2]Universität Mannheim, Germany; [3]Universität Augsburg, Germany; [4]Eberhard-Karls-Universität Tübingen, Germany

Hohe Studienmotivation ist für den akademischen Erfolg von Studierenden entscheidend (Schneider & Preckel, 2017). Die situative Erwartungs-Wert Theorie (SEVT, Eccles & Wigfield, 2020) betont die situative Natur der Lern- und Leistungsmotivation und ist eine etablierte Grundlage für Forschung im Bildungsbereich (Cook & Artino, 2016). Gemäß der Theorie sind die motivationsrelevanten Kernkomponenten – Erfolgserwartungen und subjektiver Aufgabenwert – in einen bestimmten Kontext eingebettet, an einen konkreten Zeitpunkt gebunden und werden durch die bisherige Entwicklung des Individuums beeinflusst. Der subjektive Aufgabenwert wird klassischerweise in die Facetten intrinsischer Wert, Wichtigkeit, Nützlichkeit und motivationale Kosten gegliedert, wobei umstritten ist, ob letztere als separate Motivationskomponente betrachtet werden sollten (Eccels & Wigfield, 2020). Zur Klärung solcher Fragestellungen ist es essenziell, die einzelnen Motivationskomponenten möglichst differenziert und angepasst an den jeweiligen Lebenskontext der Befragten (z.B. Schule, Hochschule) zu erfassen (Gaspard et al., 2015).

Bislang wurde kein Instrument im deutschsprachigen Raum validiert und normiert, welches eine derart lernkontextbezogene Operationalisierung von Studienmotivation ermöglicht. Schnettler et al. (2020) haben bereits eine SEVT-basierte Konzeption des subjektiven Aufgabenwerts von Studierenden und eine entsprechende Skala (Motivation: Value and Expectancies for Success scale; MoVE) entwickelt. Dabei werden bestehende Items für die individuelle Erfolgserwartung (Fleischer et al., 2019) genutzt und die zentralen Facetten des Aufgabenwertes (intrinsischer Wert, persönliche Wichtigkeit, Wichtigkeit guter Leistungen, Nützlichkeit für den Beruf, Nützlichkeit für die persönliche Entwicklung, Anstrengungskosten, emotionale Kosten, Opportunitätskosten) durch neue Subskalen gemessen. Die Skala erlaubt erstmals eine differenzierte Betrachtung der Motivationskomponenten von Studierenden, wurde aber bisher lediglich anhand einer nicht-repräsentativen Stichprobe auf faktorielle sowie konvergente und diskriminante Validität geprüft (Schnettler et al., in Vorbereitung). In der vorliegenden Studie soll die Skala durch das Überprüfen der Faktorenstruktur, internen Konsistenz, Messinvarianz sowie konvergenter und diskriminater Validität möglichst umfangreich validiert und durch die Bereitstellung von Normwerten praktisch anwendbar gemacht werden.

Zum Erhalt repräsentativer Daten wurde eine quotierte Stichprobe entsprechend der demographischen Daten der deutschen Studierendenschaft (Statistisches Bundesamt, 2021) angestrebt. Basierend auf den Variablen Geschlecht, Fachsemester, Studienrichtung und Hochschultyp wurden repräsentative Cluster gebildet, um eine verhältnismäßige Rekrutierung zu ermöglichen. Insgesamt beantworteten N = 980 Studierende den Fragebogen, wobei nicht alle Cluster vollständig gefüllt werden konnten. Die endgültige Festlegung der dafür nötigen Gewichtungsfaktoren und die finale Auswertung stehen derzeit noch aus, insofern werden im Folgenden vorläufige Ergebnisse berichtet.

Aufgrund der zahlreichen Komponenten des Aufgabenwertes empfehlen Part et al. (2020) die Faktorenstruktur mithilfe eines explorativen Strukturgleichungsmodells statt einer konfirmatorischen Faktorenanalyse zu überprüfen. Dadurch können beispielsweise Kreuzladungen der Items identifiziert werden. Die Modellierung eines B -ESEM Modells (Gu et al., 2017), in welchem die Items einer Subskala jeweils auf den spezifischen Wertdimensions-Faktoren und auf einem übergreifenden Wert-Faktor laden, lieferte gute Fit-Indizes ($c2(252) = 823.23$; $p < .001$; *RMSEA* = .050; *CFI* = .991; *SRMR* = .043).

Die internen Konsistenzen der Subskalen wurden durch die Berechnungen von Cronbachs Alpha und McDonalds Omega überprüft. Alle Werte lagen konsistent im sehr guten Bereich über .80 (Dunn & Baguley, 2014). Zwischen den Geschlechtern zeigte sich metrische Messinvarianz, was eine gruppenübergreifende Bedeutungsäquivalenz der Items nahelegt.

Zusätzliche Evidenz für die Validität lieferte die Überprüfung konvergenter und diskriminanter Validität. Motivationsförderliche Wertfacetten korrelierten erwartungsgemäß positiv mit Studienzufriedenheit (*r* zwischen .22 und .57) und Studienleistung (*r* zwischen .24 und .41) und negativ mit Studienabbruchsintentionen (*r* zwischen -.17 und -.46). Motivationale Kosten hingegen korrelierten positiv mit Studienabbruchsintentionen (*r* zwischen .12 und .38).

Die MoVE Skala macht einen vielversprechenden Gesamteindruck. Faktorielle Struktur, interne Konsistenz und Korrelationsanalysen bestätigen, dass die Skala spezifische Komponenten der Studienmotivation reliabel und valide erfasst. Sie schafft ein präzises Verständnis der Wertfacetten im Sinne der SEVT und ermöglicht differenzierte Analysen zwischen Studienmotivation und Studienerfolgsindikatoren. Sie kann sowohl für Projekte in der Hochschulforschung als auch zur individuellen Diagnostik in der Hochschulpraxis genutzt werden.

Einsatz von maschinellem Lernen zum Unterstützung der Studienauswahl für eine Metaanalyse in Bildungskontext

Olga Chernikova, Matthias Stadler, Frank Fischer
Ludwig Maximilian Universität, München

Metaanalysen tragen zu einer erhöhten Generalisierbarkeit der Ergebnisse bei, weil sie ermöglichen, die Ergebnisse einzelner Studien zu aggregieren und mögliches Rauschen durch methodologische Unterschiede zu reduzieren (z.B. Xiong et al., 2018). Jedoch führt die exponentielle Zunahme der Anzahl an Publikationen von Primärstudien in einigen Bereichen (z.B. simulationsbasiertes Lernen) in den letzten Jahren zu einer Reihe von Komplikationen bei der systematischen und vollständigen Erfassung des Forschungsstands (z.B. Ware & Mabe, 2015). Dazu gehören Fehler, systematischer Bias und erhöhte Kosten, die mit der manuellen Auswertung und Kodierung von Primärstudien verbunden sind. Dieser Beitrag beschreibt die Entwicklung, Validierung und Anwendung von maschinellen Lernalgorithmen zur Unterstützung der Klassifizierung während des Abstract Screenings im Bildungsbereich. Zentrales Ziel ist es, den Auswahlprozess durch Verkürzung der Zeit und Erhöhung von Transparenz und Objektivität zu unterstützen.

Zum Zweck des Trainings, der Validierung und des Tests verwendeten wir 3.209 Studien, die manuell als geeignet oder ungeeignet für eine Metaanalyse zu den Auswirkungen von simulationsbasiertem Lernen auf Entwicklung von komplexe kognitive Fähigkeiten markiert wurden (Autoren, 20XX).

Geeignete Studien sollten eine Intervention beschreiben, in der Simulation fürs Lernen komplexe kognitive Fähigkeiten im Hochschulkontext genutzt war und eine Kontrollbedingung beinhalten (Pre-Test oder Kontrolgruppe ohne Simulation). Von diesen 3.209 Studien wurden 7,3% als geeignet und 92,7% als nicht geeignet eingestuft. Die codierten Abstracts wurden vorverarbeitet und nach dem Zufallsprinzip in drei Teile mit einem Verhältnis von 8:1:1 aufgeteilt (Trainings-, Validierungs- und Testdatensatz). Die Vorverarbeitung umfasste das Entfernen von Füllwörtern (z.B. Artikeln, Präpositionen) und Satzzeichen, das transformieren aller Wörter in Kleinbuchstaben und in ihre Stammform („bag-of-words“ Frequenztechnik; Kwartler, 2017). Als letzten Schritt wurde für den Trainingsdatensatz die Anzahl der Merkmale mithilfe einer linearen Korrelationsanalyse zwischen den extrahierten Merkmalen und dem zur Klasse gehörenden Merkmal reduziert.

Drei verschiedene Algorithmen wurden trainiert: ein Linear Support Vector Machine (SVM)-Klassifikator (Suykens & Vandewalle, 1999), eine regularisierte logistische Regression (RLR) (Tibshirani, 1996) und ein Random Forest (Svetnik et al., 2003) zur Vorhersage von Eignung der Studien aus den Wörtern ihren Abstracts. Die, für die Textklassifikationsaufgabe als angemessen erachtete, Vorhersagegüte (mindestens 80 % Recall und 15 % Precision) wurde durch die Leistung menschlicher Codierer*innen bei derselben Aufgabe bestimmt. Unser Ziel war es, einen hohen Recall zu erreichen, aber gleichzeitig den Arbeitsaufwand für Forscher*innen zu reduzieren. Die Leistung jedes Modells wurde anhand der unsichtbaren 10 % der Daten (Testdatensatz) bewertet.

Basierend auf den Ergebnissen wurden SVM und RLR ausgewählt, um die Reihe von 2373 Studien zu klassifizieren, die in den letzten drei Jahren gefunden wurden. Die Algorithmen einigten sich auf 87% aller Abstracts. Die anderen 13% wurden von zumindest einem der Algorithmen als geeignet gekennzeichnet. Diese Abstracts wurden von erfahrenen menschlichen Bewertern begutachtet und 58 Abstracts wurden in die Volltextscreening aufgenommen. Insgesamt wurden 458 (19 % der ursprünglichen Anzahl der Abstracts) in die Volltextprüfung verschoben. Durch den Einsatz von Algorithmen konnten etwa 2/3 der für das abstrakte Screening aufgewendeten Zeit eingespart werden.

Die Ergebnisse zeigen, dass maschinelle Lernalgorithmen den Prozess der Klassifizierung von Studien für metaanalytische Zwecke effektiv unterstützen können. Die SVM- und RLR-Klassifikatoren können den Zeitaufwand für die manuelle Klassifizierung deutlich reduzieren und sind präziser als menschliche Klassifikatoren. Kontinuierliches, automatisches Pre-Screening für Metaanalysen kann von großer Bedeutung für die Verbesserung Objektivität und Generalisierbarkeit der Forschung sein. Allerdings ist zu beachten, dass das hohe Missverhältnis zwischen geeigneten und nicht geeigneten Studien die Präzision der Algorithmen einschränkt. Darüber hinaus sind die Algorithmen stark kontextspezifisch und müssen für neue Daten, die in einem anderen Kontext verwendet werden sollen, neu trainiert werden. Zusammenfassend sollte der Einsatz von maschinellen Lernalgorithmen zur Unterstützung der Forschung unbedingt weiter untersucht werden. Wir diskutieren auch weitere mögliche Verwendungsmöglichkeiten solcher Algorithmen in der Bildungsforschung (z.B. für das Design von Lernumgebungen).

„Ich kann, ich will, aber es ist anstrengend“ – Erwartungs-Wert-Kosten-Profile der Studienmotivation und ihre Zusammenhänge mit Studienzufriedenheit

Maike Trautner, Carola Grunschel
Westfälische Wilhelms-Universität Münster, Germany

Studienmotivation ist ein wichtiger Prädiktor für Studienerfolg und Wohlbefinden (Deci & Ryan, 2000; Schneider & Preckel, 2017). Das Erwartungs-Wert-Framework beschreibt Studienmotivation als Zusammenspiel aus subjektiver Erfolgserwartung, Aufgabenwert, sowie subjektiven Kosten (Eccles & Wigfield, 2020). Es ist allerdings umstritten, ob Kosten einen eigenständigen Faktor oder eine Unterfacette des Aufgabenwerts darstellen, der unabhängig vom Aufgabenwert ausgeprägt sein kann (Barron & Hulleman, 2015; Eccles & Wigfield, 2020). Die meisten bisherigen Studien schließen, dass Kosten eher einen eigenständigen Faktor darstellen, der über die anderen Wert-Facetten hinaus mit Kriterien wie Studienzufriedenheit zusammenhängt (e.g., Jiang et al., 2020; Kryshko et al., 2022). Jedoch haben die meisten dieser Studien variablenzentrierte Analysemethoden (z.B. konfirmatorische Faktorenanalysen, Regressionsanalysen) verwendet. Diese Ansätze weisen aufgrund der hoch korrelierten Wert-Facetten verschiedene Probleme auf (Wigfield & Eccles, 2020). Außerdem beleuchten sie nicht die intrapersonalen Muster bzw. Konstellationen von Erwartungs-, Wert- und Kosteneinschätzungen. Solche Analysen können außerdem Aufschluss darüber geben, welche individuellen Motivationskonstellationen mit besonders positiven oder negativen Konsequenzen einhergehen. Dadurch bieten sie differenzierte Anhaltspunkte, für welche Personen welche Maßnahmen zur Förderung von Studienmotivation besonders wertvoll sind. Außerdem können sie Hinweise liefern, ob Kosten einen eigenständigen Faktor bilden, wenn sich beispielsweise Profile mit hohem Wert und Erwartung bei gleichzeitig hohen Kosten finden (statt gleichzeitig niedrigen Kosten) oder Kosten besonders zur Differenzierung verschiedener Profile beitragen (Conley, 2012). Bisherige personenzentrierte Analysen hierzu haben Kosten allerdings oft nicht oder nur undifferenziert berücksichtigt (z.B. Conley, 2012, Hodis & Hodis, 2020, Perez et al., 2019). Die vorliegende Studie untersucht daher mittels latenter Profilanalyse (Ferguson et al., 2020) a) welche Erwartung-Wert-Kosten-Profile sich bei Studierenden zeigen, insbesondere, ob sich Evidenz für Profile mit hohem Wert und gleichzeitig hohen Kosten findet, b) ob sich diese Profile hinsichtlich ihrer Studienzufriedenheit unterscheiden und c) welche demografischen Faktoren die Profilzugehörigkeitswahrscheinlichkeit vorhersagen. Insgesamt N = 475 Medizinstudierende beantworteten Aussagen zur Erfolgserwartung (Fleischer et al., 2019, ω = .86), neun studienbezogenen Wert- und Kostenfacetten in Anlehnung an Gaspard et al. (2015, ω = .86-.92), sowie zur Studienzufriedenheit (Schiefele & Jacob-Ebbinghaus, 2006, ω = .81-.87). Die beste Profillösung auf Basis statistischer und inhaltlicher Kriterien (Marsh et al., 2004) zeigte sich bei vier distinkten latenten Profilen. Profil 1 (n = 50) zeichnete sich durch im Vergleich zu anderen Studierenden geringe Erfolgserwartung, Aufgabenwerte und hohe Kosten aus („unmotiviertes Profil“), während Profil 3 dasselbe Muster mit moderateren Kosten zeigte („moderat unmotiviertes Profil“, n = 117). Profil 2 zeigte sich als „motivational günstiges Profil“ mit hohen Erfolgserwartungen, hohem Wert und geringen Kosten (n = 164). Das vierte Profil schien „motiviert trotz Kosten“ und charakterisierte sich durch vergleichsweise hohe Erwartung, Werte und moderate Kosten (n = 144), welches am ehesten einem Profil mit hohem Wert trotz hohen Kosten nahekam.

Frauen zeigten eine höhere Wahrscheinlichkeit im „motivational günstigen Profil“ zu sein im Vergleich zu allen anderen Profilen. Je weiter fortgeschritten Lernende in ihrem Studium waren, desto weniger wahrscheinlicher fielen sie in das motivierte Profil (Profil 2) verglichen mit allen anderen Profilen. Wie zu erwarten zeigte das motivational günstige Profil höhere Studienzufriedenheiten bzgl. Inhalten, Bedingungen, sowie dem Umgang mit Belastungen im Vergleich zu allen anderen Profilen.

Die ermittelten Profile ähneln bisherigen personenzentrierten Studien, in denen besonders Kosten zwischen verschiedenen Motivationsprofilen differenzierten (Conley, 2012) und Studien mit ähnlich großen Stichproben und globalerer Erfassung von Wert und Kosten (Hodis & Hodis, 2020, Perez et al., 2019). Die differenzierte Erfassung der Wert- und Kostenfacetten in der aktuellen Studie lieferte keinen Mehrwert zur Profildifferenzierung gegenüber Profilen auf Basis der globalen Konstrukte (Hodis & Hodis, 2020). Die Ergebnisse geben Anhaltspunkte, dass besonders bei Lernenden mit geringem Aufgabenwert und Erfolgserwartungen bei gleichzeitig hohen Kosten in höheren Semestern Maßnahmen zur Motivationsförderung angeraten sind, um Unzufriedenheit im Studium zu reduzieren.

B-10: Akademische Professionalisierung

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S05 R03 H20
Chair der Sitzung: Laura Dörrenbächer-Ulrich

Entwicklung und Validierung eines Strategiewissenstests zum selbstregulierten Lernen bei Studierenden

Laura Dörrenbächer-Ulrich, Jörn R. Sparfeldt, Franziska Perels
Universität des Saarlandes, Germany

Selbstreguliertes Lernen (SRL) als zielgerichteter Prozess, der kognitive, metakognitive und motivationale Komponenten umfasst (Perels et al., 2020), ist hoch relevant für akademische Leistungen sowie das Wohlbefinden Studierender (DiBenedetto & Bembenutty, 2013). Die vorliegenden Methoden und Instrumente zur Erfassung von SRL weisen je nach Anwendungskontext spezifische Vor- und Nachteile auf (Rovers et al., 2019). Mit sog. Strategiewissenstests können die Komponenten selbstregulativen Wissens als Kompetenz gemessen werden (Wirth & Leutner, 2008) und so das Wissen über die Passung verschiedener Strategien zu bestimmten exemplarischen Problemsituationen (Vignetten) erfassen. Dabei sollen die Proband*innen die Nützlichkeit verschiedener SRL-Strategien zur Lösung eines Lernproblems bewerten (unabhängig ihrer eigenen Nutzungswahrscheinlichkeit). Obwohl bereits vereinzelte SRL-Strategiewissenstests existieren (z.B. Händel et al., 2013), sind Tests für Studierende selten und beziehen sich oft nur auf ausgewählte SRL-Komponenten. Die vorliegende Studie zielt daher auf die Entwicklung und Validierung eines SRL-Strategiewissenstests für Studierende, der alle drei SRL-Komponenten im Rahmen des SRL-Prozesses (Planung, Handlung, Reflexion; Zimmerman, 2000) abdeckt.

Studie 1 stellt die theoretische Entwicklung des Instruments vor und berichtet die Ergebnisse einer Expert*innenbefragung. Es wurden insgesamt sieben Lernszenarien entwickelt: pro SRL-Phase jeweils ein metakognitives und ein motivationales Szenario sowie ein kognitives Szenario für die Handlungsphase. Zu jedem Szenario wurden (im Hinblick auf das konkrete Szenario) drei nützliche und drei weniger nützliche Strategien präsentiert. Siebzehn Expert*innen sollten die Szenarien den SRL-Komponenten und -Prozessen zuordnen. Außerdem sollten Strategien zur Lösung des Lernproblems vorgeschlagen werden. Die Zuordnungen und vorgeschlagenen Strategien stimmten größtenteils mit unseren a priori Einschätzungen überein. Darüber hinaus sollten die Expert*innen die Nützlichkeit von sechs Strategien zur Lösung des beschriebenen Lernproblems auf einer vierstufigen Skala bewerten. Die Bewertung stimmte mit unserer Einschätzung ebenfalls überein und bestätigte damit die ausgewählten Szenarien und Strategien.

Studie 2 stellt eine Pilotierungsstudie mit n = 143 Studierenden dar. Die Studierenden sollten die jeweils sechs Strategien auf einer vierstufigen Skala hinsichtlich ihrer Nützlichkeit bewerten. Durch Paarvergleiche wurden Scores berechnet, deren Mittelwerte nicht auf Deckeneffekte hindeuteten. Die Schwierigkeiten zeigten eine angemessene Varianz und die Reliabilitätshinweise (Cronbachs α) (sehr) gute Werte. Die Komponenten korrelierten untereinander sowie mit einem SRL-Selbstberichtsfragebogen in niedriger bis mittlerer Größe. Zur Studienleistung zeigte sich kein signifikanter Zusammenhang. Als Konsequenz wurde die Scoringmethode angepasst.

Studie 3 untersuchte den Zusammenhang des Strategiewissenstests zu einem SRL-Fragebogen sowie zu SRL-Mikroanalysen (Strategieerfassung während einer Lernsituation) an einer Stichprobe

von *n* = 99 Studierenden. Wie in der Pilotierungsstudie zeigten sich keine Deckeneffekte, angemessene Schwierigkeiten sowie (sehr) gute Reliabilitätshinweise. Innerhalb des Instrumentes ergaben sich mittlere bis hohe Korrelationen zwischen den Komponenten. Zum Fragebogen sowie zu den Mikroanalysen resultierten niedrige bis mittlere Korrelationen. Mit der Abiturnote sowie einem Leistungstest korrelierte der Strategiewissenstest nicht bedeutsam.

Studie 4 analysierte die Zusammenhänge des Strategiewissens zu weiteren relevanten Konstrukten für Studierende sowie die Stabilität des Strategiewissens. Insgesamt nahmen *n* = 207 Studierende zum ersten Messzeitpunkt und *n* = 106 Studierenden nach acht Wochen zum zweiten Messzeitpunkt teil. Auch hier zeigten sich keine Deckeneffekte, angemessene Schwierigkeiten sowie (sehr) gute Reliabilitätshinweise. Eine konfirmatorische Faktorenanalyse ergab eine gute Modellanpassung. Innerhalb des Strategiewissenstests zeigten sich erneut mittlere bis hohe Korrelationen – neben mittleren Korrelationen zum SRL-Fragebogen. Es fanden sich außerdem mittlere Korrelationen zur SRL-Nützlichkeitseinschätzung, Studienzufriedenheit, Selbstwirksamkeit und zum Selbstkonzept. Mit dem Wohlbefinden und der Testängstlichkeit ergaben sich keine bedeutsamen Zusammenhänge. Zur Abiturnote wurden schwache signifikante Korrelationen gefunden. Zudem konnte eine hohe Stabilität die SRL-Komponenten dokumentiert werden.

Die vier Studien zeigen insgesamt, dass der neu entwickelte Strategiewissenstest gut geeignet ist, um das Wissen über SRL-Strategien bei Studierenden zuverlässig zu erfassen. Auch zeigen sich erwartbare Zusammenhänge zu anderen SRL-Maßen bzw. studienrelevanten Konstrukten. Der (größtenteils) fehlende Zusammenhang zu Leistungsvariablen muss diskutiert werden.

Academic Self-Concept and Goal Orientations among Adolescents from different Learning Environments

Olga Steinberg, Stefan Kulakow, Diana Raufelder
Universität Greifswald, Germany

Theoretical background

Adolescence presents a challenging developmental period for all students, and is crucial for the declines in academic self-concept (e.g., Nagy et al., 2010, Pesu, et al., 2016) and academic motivation (e.g., Gnambs & Hanfstingl, 2016) that goes together with the downward trend in performance (e.g., Wijsman, 2015). One of the main reasons for this decline lies in the institutional transition from primary to secondary school, and a mismatch between students' developmental needs and increased demands from teachers, particularly in traditional teacher-directed learning (TDL) environments. Less is known about students' academic self-concepts, academic motivation and performance in student-directed learning (SDL) environments. Therefore, we explore the differences in the association of domain-specific academic self-concepts and performance, mediated by goal orientations, in a traditional TDL school setting as opposed to SDL, based on the Reciprocal Effects Model (Marsh,1990; Marsh & Craven, 2006; Marsh & Martin, 2011).

Research questions

Based on the existing research, we hypothesize that:

- Students from SDL report higher verbal and mathematic self-concepts, and mastery goal orientation in comparison to students from TDL. In contrast, students from TDL would report higher performance goal orientations.

- The learning environment moderates the longitudinal association of mathematic self-concept and grades, mediated by goal orientations.

- The learning environment moderates the longitudinal association of verbal self-concept and grades, mediated by goal orientations.

Method

1,153 students from Mecklenburg-Vorpommern, Germany (Mage = 13.97; SD = 1.37, 49% girls) from grades 7–10 participated in the study. These students attended 57 classrooms in six secondary schools in Germany. The students' self-report data was collected twice. The students were surveyed during the winter term 2015/ 2016 (T1) and 0.5 years later (T2). From the six schools, that comprise the sample, three used SDL, based on competency grids, whereas the other three followed a TDL approach. As self-concept is domain-specific, we addressed self-concepts for mathematics and German by two scales from the PISA 2000 questionnaire (Artelt et al., 2004). We used the scales developed by Spinath et al. (2012) to address goal orientations. Finally, we collected data on grades and gender. To examine the hypotheses, we used a multigroup cross-lagged panel analyses with additional indirect effects, conducted in Mplus 8.5 (Muthén & Muthén, 1998–2020).

Results

The results do not fully support our hypotheses as there are hardly any differences between SDL and TDL. For example, contrary to our expectations, the mean differences between SDL and TDL groups were significant for goal orientations only. TDL students reported higher values in performance-approach as well as performance-avoidance goals than students from SDL. In contrast to hypotheses II and III, no moderation effect for the learning environments could have been found. In other words, the learning environment seems to play a subordinate role. Further details and practical implications for the findings will be discussed at the conference.

Der Zusammenhang von Fehler-Überzeugungen Studierender mit deren akademischer Risikobereitschaft

Vanessa Hübner, Maximilian Pfost
Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund

Der Umgang mit Fehlern in Bildungskontexten gewinnt seit den 1990er-Jahren sowohl im Rahmen der Schul- als auch der Hochschulbildung Relevanz (Chott, 1999; Soutter & Clark, 2021). Besonderes Forschungsinteresse im Hochschulkontext besteht an *Fehler-Überzeugungen*. Überzeugungen beinhalten individuell subjektive Bewertungen eines Gegenstands hinsichtlich Gefühlen, Wissen und potenziellen Handlungsoptionen. Vorangehende Studien zeigen, dass diese positiv mit akademischem Erfolg (Leighton et al., 2018), sowie der Anwendung von Lernstrategien (Reindl et al., 2020) Studierender zusammenhängen. Unklar ist, ob die Fehler-Überzeugungen Studierender mit deren Bereitschaft zusammenhängen, Lernsituationen zu nutzen, in denen es wahrscheinlich ist

Fehler zu machen. Diese Bereitschaft kann auch als akademische Risikobereitschaft bezeichnet werden (Beghetto, 2009; Clifford, 1991). Studierende, die akademische Risiken eingehen beteiligen sich trotz Unsicherheit am universitären Geschehen. Dabei sind sie sich der Möglichkeit bewusst, Fehler zu machen und von Dozierenden beziehungsweise Mitstudierenden als weniger kompetent wahrgenommen zu werden. Die Forschung zeigt, dass akademische Risikobereitschaft dem Lernerfolg Studierender zuträglich ist (Krochmal & Roth, 2017). Dennoch neigen Studierende dazu, akademische Risiken zu vermeiden (Teagarden et al., 2018). Der Mechanismus, der diesem Phänomen zugrunde liegt ist weitgehend unerklärt, doch die Literatur legt nahe, dass positive Überzeugungen zu stärkerer und häufigerer Beteiligung an Verhaltensweisen führen, bei denen der Überzeugungsgegenstand vorherrschend ist (Ajzen & Fishbein, 1975). Insofern besteht ein Interesse an der Klärung der Frage, ob die Fehler-Überzeugungen der Studierenden einen Einfluss auf die Bereitschaft haben, akademische Risiken einzugehen und fehleranfällige Lernsituationen als potenzielle Lerngelegenheiten zu nutzen.

Fragestellung

Kann die akademische Risikobereitschaft Studierender durch deren Fehler-Überzeugungen vorhergesagt werden?

Methode

Fehler-Überzeugungen wurden an einer Stichprobe von *N* = 159 Studierenden der Human-, Sozial-, sowie Geistes- und Kulturwissenschaften (weiblich = 115, männlich = 27, *M* = 24.92 Jahre) unter Berücksichtigung von Affekt (z.B. „Wenn ich im Seminar einen Fehler mache, dann ist mir das peinlich“), Kognition (z.B. „Gut zu studieren bedeutet, wenige Fehler zu machen“) und Verhalten (z.B. „Wenn ich einen Fehler gemacht habe, sollte ich ihn mit meinen Mitstudierenden besprechen“) erfasst.

Akademische Risikobereitschaft wurde zweidimensional, einerseits in Hinblick auf die gesamte Seminargruppe (einschließlich der/des Dozierenden; z.B. „Ideen mit dem Plenum teilen, auch wenn sie noch unvollständig sind“) sowie andererseits exklusiv vor Mitstudierenden (z.B. „Mitstudierenden zu Seminarinhalten, die mir unklar geblieben sind Fragen stellen“) erfasst.

Zusammenhänge wurden mithilfe von schrittweiser linearer Regression geprüft.

Ergebnisse

Die Auswertungen zeigen, dass die Fehler-Überzeugungen Studierender deren Bereitschaft vorhersagen, akademische Risiken vor der gesamten Seminargruppe einzugehen. Studierende, die nach Fehlern weniger negativen Affekt zeigen gehen signifikant mehr akademische Risiken ein (β = .43), während Studierende, die Fehlern behavioral zugeneigt sind signifikant weniger akademische Risiken eingehen (β = -.20). Die kognitive Fähigkeit, Fehler als Lerngelegenheit zu betrachten hat keinen Einfluss auf die akademische Risikobereitschaft der Studierenden. Weiterhin wurde kein Einfluss von Fehler-Überzeugungen auf die Bereitschaft Studierender gefunden, akademische Risiken vor ihren Mitstudierenden einzugehen.

Diskussion

Die Ergebnisse der Untersuchung tragen zum Verständnis der Rolle von Fehler-Überzeugungen Studierender für Bildungsdynamiken und Bildungsprozesse bei. Während das rationale Wissen um die Nützlichkeit von Fehlern scheinbar nicht ausreicht um akademische Risikobereitschaft im Rahmen universitärer Seminare vorherzusagen, besitzt die Abwesenheit von negativem Affekt Relevanz. Diese beiden Ergebnisse erklären, warum Studierende akademische Risiken vermeiden, obwohl sie diese als sinnvoll erachten (Ravert & Schneller, 2019). Der negative Zusammenhang zwischen der behavioralen Fehlerzugewandtheit und akademischer Risikobereitschaft deutet auf das Motiv der Arbeitsvermeidung hin. Die fehlenden Zusammenhänge zwischen Fehler-Überzeugungen und akademischer Risikobereitschaft vor Mitstudierenden können darauf hindeuten, dass Fehler in diesen Situationen keine subjektive Bedrohung für Studierende darstellen. Lerngruppen finden sich zumeist aufgrund von Sympathie oder gemeinsamen Interessen zusammen, was bei der Entscheidung akademische Risiken einzugehen eine wichtigere Rolle spielen mag.

Effekte einer Lehrkräftefortbildung auf die professionellen Kompetenzen im selbstregulierten Lernen: Befunde aus einer Interventionsstudie

Carmen Nadja Hirt, Yves Karlen, Johannes Jud, Amina Rosenthal, Tabea Eberli
Pädagogische Hochschule FHNW, Schweiz

Die professionellen Kompetenzen von Lehrkräften im selbstregulierten Lernen (SRL) beeinflussen ihr Unterrichtshandeln (Karlen et al., 2020). Die Forschung weist jedoch darauf hin, dass die Lehrkräfte teils nicht nur über geringe professionellen Kompetenzen im SRL verfügen (z. B. Lawson et al., 2019; Porter & Peters-Burton, 2021), sondern auch die Förderung des SRL eine Herausforderung für sie darstellt, was sich in geringfügigen Fördertätigkeiten im Unterricht niederschlägt (Dignath & Veenman, 2021). Folglich ist die Entwicklung der professionellen Kompetenzen von Lehrkräften im SRL eine wichtige Voraussetzung, damit das SRL vermehrt in den Unterricht integriert wird.

Bisherige Evaluationen von Fortbildungen im SRL fokussierten insbesondere Effekte auf die Förderpraxis von Lehrkräften. Während einige Untersuchungen positive Effekte auf die Förderpraxis nachweisen konnten (Kramarski & Revach, 2009; Werth et al., 2012), zeigten andere Analysen eingeschränkte (lediglich bei Lehrkräften mit geringem Vorwissen, Venitz & Perels, 2018) oder keine Effekte (Perels et al., 2009). Basierend auf dem integrativen Modell der professionellen Kompetenzen von Lehrkräften im SRL (Karlen et al., 2020), sollten Fortbildungen über die Effekte auf die Förderpraxis der Lehrkräfte hinausgehen und somit die Entwicklung der Kompetenzen von Lehrkräften als SRL-Vermittler*innen sowie als selbstregulierte Lernende fokussieren (doppeltes Kompetenzprofil).

Bisherige Untersuchungen zeigen positive Effekte von Interventionen auf die Selbstwirksamkeit (Dignath, 2021) und das Wissen (Perels et al., 2009). Die Untersuchung von Dignath (2021) konnte hingegen keine Effekte auf das Wissen sowie auf die Überzeugungen finden. Mit Bezug zu den Kompetenzen der Lehrkräfte als selbstregulierte Lernende wiesen Perels et al. (2009) einen positiven Effekt auf die selbstberichtete Selbstregulierung nach. Gesamthaft ist die Befundlage zu Lehrkräftefortbildungen im SRL noch undeutlich. Zudem werden insbesondere Untersuchungen benötigt, die das doppelte Kompetenzprofil von Lehrkräften berücksichtigen.

Die vorliegende Untersuchung analysiert die Effekte einer mehrtätigen Lehrkräftefortbildung im SRL (vier Tage verteilt über sechs Monate in den Jahren 2021/22) auf die Kompetenzen der Lehrkräfte (1) als selbstregulierte Lernende und (2) als SRL-Vermittler*innen. Im Rahmen der Interventionsstudie mit Prä-Post- und Kontrollgruppendesign wurden N = 54 Lehrkräfte (n_{EG} = 31; n_{KG} = 23) mit einer durchschnittlichen Berufserfahrung von 16.48 Jahren (SD = 10.13) befragt. Die Lehrkräfte waren zu 61.1% weiblich und im Durschnitt M = 34.56 Jahre (SD = 8.95) alt.

Mittels Selbstberichten und Wissenstests wurde das eigene SRL der Lehrkräfte (metakognitive, motivationale Regulationskompetenzen) und deren Kompetenzen als SRL-Vermittler*innen (PCK-SRL, Überzeugung zur Bedeutung der Fremd-/Selbstregulation, Selbstwirksamkeit Förderung, Werte [Nutzen, Interesse, persönliche Bedeutung], Kosten [Effort-, Opportunitätskosten]) erfasst (α_{t1} = .69 - .94; α_{t2} = .60 - .94). Die Förderpraxis wurde mit einer Skala zu SRL-Unterrichtstätigkeiten erhoben (α_{t1} = .75, α_{t2} = .74). Anhand von mixed ANOVAs wurden mögliche Interaktionseffekte (Zeit*Gruppe) untersucht.

Hinsichtlich der Kompetenzen der Lehrkräfte als selbstregulierte Lernende ergeben die Analysen einzig einen signifikanten Interaktionseffekt für die motivationale Regulationskompetenz, $F(1, 48)$ = 4.419, p = .041, partielles η^2 = .084. Die Experimentalgruppe wies nach der Intervention höhere Werte auf als die Kontrollgruppe. Hinsichtlich der Kompetenzen der Lehrkräfte als SRL-Vermittler*innen ergaben sich bedeutende Interaktionseffekte für die Wertkomponenten Interesse, $F(1, 47)$ = 7.118, p = .010, partielles η^2 = .132, und persönliche Bedeutung, $F(1, 45)$ = 4.259, p = .045, partielles η^2 = .086, wobei auch hier die Experimentalgruppe nach der Intervention jeweils höhere Werte aufwies als die Kontrollgruppe. Für die weiteren untersuchten Kompetenzbereiche konnten keine Interaktionseffekte gefunden werden.

Gesamthaft weisen die Ergebnisse somit lediglich auf Effekte bezüglich motivationaler Kompetenzen der Lehrkräfte hin. Zum einen gilt es zu diskutieren, wie eine Lehrkräftefortbildung im Bereich SRL zu konzipieren ist, sodass sich die erwünschten Effekte zeigen. Zum anderen zeigt sich, dass es im Bereich von Lehrkräftefortbildungen kein «One size fits all» gibt. Weitere differenzierte Analysen sind notwendig, um herauszufinden, wer von einer Lehrkräftefortbildung profitiert.

Symposien

B-11: Zur quantitativen Vermessung kultureller Bildung – methodologische Reflexionen zu Datenverfügbarkeit und Operationalisierung am Beispiel aktueller Forschungsansätze

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* S07 S00 D07

Zur quantitativen Vermessung kultureller Bildung – methodologische Reflexionen zu Datenverfügbarkeit und Operationalisierung am Beispiel aktueller Forschungsansätze

Chair(s): **Jannis Burkhard** (DIPF, Deutschland), **Stefan Kühne** (DIPF)
*Diskutant*in(nen):* **Katrin Rakoczy** (Justus-Liebig-Universität Gießen)

Aktuell wird dem Feld der kulturellen Bildung oft ein bildungspolitischer Bedeutungszuwachs attestiert. Dieser wird zum einen häufig daran festgemacht, dass die Kultusministerkonferenz im Moment ihre Empfehlungen zu kultureller Bildung (KMK, 2013) überarbeitet und zum zweiten oftmals daran, dass das Bundesministerium für Bildung und Forschung in den letzten Jahren Forschung zur kulturellen Bildung verstärkt fördert. Lange Zeit kamen bei diesen Forschungsaktivitäten überwiegend qualitative empirische Methoden zum Einsatz (Jörissen & Liebau, 2013). Seit einigen Jahren gibt es jedoch einen Trend hin zu mehr quantitativer Forschung (Reinwand-Weiss, 2020b), durch den sich neue methodologische Herausforderungen ergeben, die dieses Symposium adressiert.

Eine dieser Herausforderungen resultiert aus der Weite des Konzepts der kulturellen Bildung. Neben den aus europäischer Sicht klassischen Kunstsparten schließt kulturelle Bildung als Dachbegriff auch soziokulturelle Praktiken ein (Reinwand-Weiss, 2013; Liebau, 2018; s. auch Reckwitz, 2010). Um quantitative Forschung betreiben zu können, muss daher die Frage geklärt werden, welche Facetten kultureller Bildung sinnvollerweise operationalisiert werden sollten.

Die begrenzte Verfügbarkeit von Daten schränkt die Möglichkeiten für quantitativ-empirische Analysen weiter ein. Beispielsweise entschloss sich die Autorengruppe Bildungsberichterstattung (2012) aufgrund des Mangels an verfügbaren Daten dazu, neue Erhebungen durchzuführen. Seitdem haben Sekundärdatenanalysen in der Bildungsforschung weiter an Bedeutung gewonnen (DGfE, GEBF & GFD, 2020). Doch in großen Erhebungen im Bildungsbereich wird kulturelle Bildung weiterhin oft nur am Rande adressiert, sodass der Sekundärdatenbestand überschaubar ausfällt.

Durch die verstärkte Verwendung von quantitativen Methoden stellt sich außerdem die Frage nach geeigneten Erhebungsinstrumenten. Grundsätzlich sind die Anforderungen an diese davon abhängig, ob die Angebots-, Prozess- oder Ergebnisdimension (vgl. Klieme et al., 2007) operationalisiert werden soll. Zu klären ist an dieser Stelle das Verhältnis zwischen kultureller Bildung und kulturellem Kapital (Bourdieu, 2012; Goßmann, 2018).

Das Symposium nähert sich den skizzierten Fragen folgendermaßen: Beitrag 1 zeigt das Spektrum auf, in dem man sich bei der Operationalisierung kultureller Bildung bewegt. Es wird der Frage nachgegangen, welche Teilbereich kultureller Bildung sinnvollerweise gemessen werden sollten. Beitrag 2 gibt einen Überblick darüber, welche Analysemöglichkeiten zu kultureller Bildung in aktuellen Sekundärdatenbeständen existieren. Zusätzlich werden Regionalbezüge der Sekundärdaten hervorgehoben und diskutiert. Beitrag 3 stellt ein Erhebungsinstrument und damit eine Operationalisierung für die Ergebnisdimension kultureller Bildung vor. Dabei wird mit Mitteln der Item-Response-Theory eine kompetenzdiagnostische Herangehensweise verfolgt. Beitrag 4 untersucht

das Verhältnis von kultureller Bildung und kulturellem Kapital. Hierfür wird eine Operationalisierung für kulturelle Bildung verwendet, die sich an der Prozessdimension des Begriffs orientiert. Im Symposium sind sowohl erziehungswissenschaftliche, musikpädagogische als auch soziologische Perspektiven vertreten.

Beiträge des Symposiums

Analysen zum Nutzungsverhalten kultureller Angebote von Kindern und Jugendlichen unter Berücksichtigung theoretischer Differenzkategorien

Sonja Nonte[1], Andreas Lehmann-Wermser[2], Julius Kopp[2]
[1]Universität Osnabrück, [2]Hochschule für Musik, Theater und Medien Hannover

Theoretischer Hintergrund

Kulturelle Bildung ist nach allgemeiner Auffassung Teil einer umfassenden Grundbildung und dient zum einen der aktiven Beteiligung am kulturellen Leben und ist zum anderen der Schlüssel für die kulturelle und gesellschaftliche Teilhabe (Ermert, 2009; Weishaupt & Zimmer, 2013; Deutsche UNESCO-Kommission, 2022). Kulturelle Bildung hat in Deutschland eine lange Tradition und wird vor allem von Kunst- und Kultureinrichtungen, Kindertagesstätten, Schulen, Universitäten, außerschulischen Einrichtungen sowie von den Medien angeboten (ebd.). Was genau der Begriff Kulturelle Bildung umfasst, ist dabei nicht eindeutig geklärt. Erscheint in Europa der Begriff arts education relativ unumstritten, so ist dies international nicht zwangsläufig gegeben. So erscheint in einigen Regionen der Erde das Stelzenlaufen und Flechten bedeutsamer als die traditionellen westlichen Kunstsparten Theater, Bildende Kunst und Musik (vgl. Fuchs, 2008). Eine Systematisierung kultureller Angebote erfolgt dabei häufig anhand der Unterteilung in formale, non-formale und informelle Organisationsformen (Thole & Höblich, 2014). Des Weiteren lässt sich zwischen rezeptiver Beschäftigung (z. B. dem Besuch eines Konzerts) einerseits und produktiver Beschäftigung (z. B. der Mitwirkung in einer Theateraufführung) andererseits unterscheiden (Kröner, 2013; Reinwand-Weiss, 2020a). Eine weitere Differenzierung wird im englischsprachigen Raum anhand der Kategorien highbrow und lowbrow culture vorgenommen. Dabei findet eine Wertung der Hochkultur (highbrow) als „abgehoben und elitär“ (Hoppe, 2011, S. 170) statt, die im Gegensatz zur volksnahen Massenkultur (lowbrow) steht. Hier wird deutlich, dass ein enger Zusammenhang zwischen der Teilnahme an (hoch)kulturellen Bildungsangeboten mit der sozialen Herkunft besteht, was bereits vielfach empirisch belegt werden konnte (u.a. Huth & Weishaupt, 2009; Kröner et al., 2008).

Fragestellung

Die Auseinandersetzung mit der Frage, inwieweit die beschriebenen Kategorisierungen sich tatsächlich auch empirisch, insbesondere auch im Kindes- und Jugendalter abbilden lassen, steht bislang noch weitgehend aus. Entsprechend soll der Beitrag dieses Desiderat aufgreifen und die theoretisch beschriebenen Differenzkategorien im Kontext von Typenbildungen (personenzentrierter Ansatz) sowie modellbasiert (variablenzentrierter Ansatz) prüfen und gegenüberstellen (Wang & Wang, 2020).

Methode

Die Grundlage bilden Daten aus zwei Studien: im Kontext der ersten Studie SIGrun (Studie zum Instrumentalunterricht an Grundschulen, gefördert vom BMBF, Lehmann-Wermser et al., 2014) wurden u.a. n=1219 Viertklässler*innen zu ihrer Teilnahme an kulturellen Aktivitäten befragt. Die Indikatoren stammen dabei sowohl aus dem hochkulturellen Bereich (u.a. Theater, Oper, Kunstausstellung) als auch dem populärem Kulturbereich (u.a. Zoo, Zirkus, Kino) und wurden anhand einer vierstufigen Likert-Skala (1="nie oder fast nie" bis 4="mehr als 4-mal im Jahr") beantwortet. Im Kontext der zweiten Studie PReTuS (Periphere Regionen, Teilhabe und Schule, gefördert vom BMBF, Brauns et al., 2022) wurden n=520 Kinder und Jugendliche der Jahrgangsstufen 4 bis 13 zu ihren jeweiligen kulturellen Aktivitäten befragt.

Ergebnisse und ihre Bedeutung

Zur ersten Studie: Die Modellierung eines Gesamtmodells kulturelle Bildung mit Subfacetten (rezeptiv/produktiv sowie Hochkultur und Massenkultur) zeigte im Kontext von CFAs unter Berücksichtigung multipel imputierter Daten (m=10) gute Anpassungswerte (X2=.116.37, df=31, p≤.001, RMSEA=.05, CFI=.96, WRMR=.95). Erwartungswidrig ließen sich hingegen auf der Basis von LCAs keine Typen mit distinkten Aktivitätsmustern identifizieren. Es zeigten sich lediglich lineare Unterschiede in der Nutzungshäufigkeit. Möglicherweise mag dies auf eine fehlende Differenziertheit bei der Erfassung kultureller Praxen im Kindes- und Jugendalter zurückzuführen sein, die nicht zuletzt durch die zunehmende Digitalisierung und Ausdifferenzierung der Lebenswelten von Kindern und Jugendlichen an Vielfalt und Diversität gewonnen haben. Entsprechend kamen im Kontext der zweiten Studie eine Vielzahl an Indikatoren zur Erfassung produktiver und rezeptiver Aktivitätsformen zum Einsatz, die neben der Abfrage klassischer hochkultureller Aktivitäten auch die Nutzung digitaler Angebote beinhaltete (Apps, Tutorials etc.). Die Daten werden derzeit ausgewertet. Entsprechend sollen die Ergebnisse im Kontext des Symposiums vorgestellt und diskutiert werden. Unter Einbezug der weiteren Beiträge wird eruiert, inwiefern die quantitative Vermessung bestimmter Facetten kultureller Bildung die Lebensrealität von Kindern und Jugendlichen adäquat abbildet und wie geeignete Operationalisierungen dafür beschaffen sein müssen.

Sekundärdatenanalyse und kulturelle Bildung. Eine Bestandsaufnahme der aktuellen Analyse- und Operationalisierungsmöglichkeiten unter Berücksichtigung räumlicher Disparitäten

Lea Fobel, Nina Kolleck
Universität Leipzig

Theoretischer Hintergrund

Der Aufschwung kultureller Bildungsforschung gründet insbesondere auch in der Annahme, dass kulturelle Bildung einen wertvollen Beitrag leistet, um aktuellen gesellschaftlichen Herausforderungen nachhaltig und einschlägig zu begegnen. Aktuelle Forschung zeigt, dass kulturelle Bildung politisches Engagement positiv beeinflusst (Fobel & Kolleck, 2021) und die persönliche Entwicklung sowie kritisches Denken fördert (Fietz, 2009; Dumitru, 2019).

Bereits bestehende Sekundärdaten für die eigene Forschung zu nutzen bietet grundsätzlich viele Vorteile sowohl für die Forschenden (z.B. Zeit, Kosten) als auch für die Beforschten (z.B. Zeit). Durch die Analyse von Sekundärdaten wird eine nachhaltige Nutzung von Daten gefördert, da Daten genutzt

werden, die der Wissenschaft zur Verfügung gestellt und damit langfristig von Dritten genutzt werden können.

Bisher ist jedoch wenig transparent, welche Datenbestände zum Thema kulturelle Bildung existieren und welche Informationen (z.B. Formalitätsniveau, regionaler Bezug) diese Daten enthalten. Des Weiteren steht das Forschungsfeld vor der Herausforderung, keine einheitliche Begriffsdefinition vorzuweisen (Reckwitz, 2010), sondern vielmehr entlang unterschiedlicher Forschungsperspektive auf diverse Verständnisse zurückzugreifen. Entsprechend unterscheidet sich auch der Umfang der Datenverfügbarkeit je nach Forschungsinteresse und Untersuchungsgegenstand (Büdel et al., 2022; Autorengruppe Bildungsberichterstattung, 2012).

Fragestellung

Der Beitrag stellt Ergebnisse eines Reviews des Sekundärdatenbestands zu kultureller Bildung in Deutschland vor und zeigt die Potentiale und Lücken quantitativer Sekundärforschung im Feld auf. Insbesondere werden Antworten auf folgende Fragen gesucht:

- Welche Aussagen können anhand der Daten zu kultureller Bildung in Deutschland über welchen Zeitraum getroffen werden?

- Welche Grundgesamtheit wird von Daten zu kultureller Bildung in Deutschland abgedeckt? Über welche Personen und Akteure können keine Aussagen getroffen werden?

- Welche Facetten kultureller Bildung werden nicht abgedeckt?

Methode

Für diesen Beitrag wurde ein Review des quantitativen Sekundärdatenbestands mit Erhebungen seit 2000 durchgeführt. Items zu kultureller Bildung wurden entsprechend eines sehr breiten Begriffsverständnisses aufgenommen, um dem breiten Verständnis vieler Forschungsprojekte gerecht zu werden (Büdel et al., 2022). Befragungen ohne Repräsentativitätsanspruch sowie administrative Datenquellen wurden nicht aufgenommen.

Ergebnisse und deren Bedeutung

Erste Ergebnisse des Reviews weisen auf Herausforderungen hin, die unter anderem aus den unterschiedlichen Operationalisierungen kultureller Bildung resultieren. Viele quantitative empirischen Studien greifen auf einen Kulturbegriff zurück, der ausschließlich musikalische sowie künstlerische Aktivitäten fokussiert.

In den großen sozialwissenschaftlichen Studien wird kulturelle Bildung wenig differenziert und häufig lediglich als kulturelles Kapital (Bourdieu, 1982) erfasst. Zusätzlich können in einigen Studien Informationen zu kultureller Bildung aus Fragen zu Freizeitaktivitäten oder Fertigkeiten abgeleitet werden. Weitere Herausforderungen ergeben sich aus der Tatsache, dass die Items nicht regelmäßig über die Befragungswellen hinweg erhoben wurden.

Räumliche Bezüge werden in diesen großen Studien häufig über die Kreiskennziffer hergestellt. Die Matchingoptionen sind jedoch eingeschränkt, da die Forschungsinstitute in der Regel umfangreiche Datenschutzmaßnahmen vornehmen, die eine intensive Verknüpfung mit Regionaldaten erschwert. Dennoch bietet das Matching über die Kreiskennziffer den Vorteil, dass jegliche Informationen dem

Datensatz hinzugespielt werden können, die auf dieser Ebene vorliegen. Viele andere Studien beschränken sich auch aus Datenschutzgründen auf die Erhebung einzelner Regionalvariablen, die entweder durch das Forschungsinstitut (z.B. Gemeindegrößenklassen) oder subjektiv durch Befragung der Interviewten (z.B. Bevölkerungsgröße des Wohnorts) dem Datensatz zugeführt werden. Dies erleichtert einerseits die Herausforderungen hinsichtlich des Datenschutzes, limitiert jedoch gleichzeitig stark die regionalen Analysemöglichkeiten.

Die Daten aus kulturwissenschaftlichen Studien (Keuchel, 2006; Keuchel & Wiesand, 2008) bieten prinzipiell vielfältige deskriptive Analysemöglichkeiten, die eine detaillierte Darstellung zu kulturellen Interessen, Partizipations- und Teilhabemuster ermöglichen und einen Einblick in die Zufriedenheit mit kulturellen Angeboten geben. Dennoch sind diese Daten nur eingeschränkt repräsentativ und aufgrund geringer Ressourcen nicht ohne Weiteres für die Wissenschaft zugänglich.

Insgesamt zeigen die Analysen, dass die verfügbaren Datensätze kulturelle Bildung nur sehr eingeschränkt abbilden und auch kleinräumige Bezüge nur sehr begrenzt in die Auswertungen einbezogen werden können. Auf der Basis dieser Erkenntnisse werden im Beitrag Chancen und Möglichkeiten von Sekundärdatenanalysen im Bereich kultureller Bildung diskutiert.

Rezeptive Bildkompetenz in der Bevölkerung. Entwicklung und Validierung eines Messinstruments für rezeptive künstlerische Fähigkeiten

Holger Lübbe, Gunnar Otte
Johannes Gutenberg-Universität Mainz

Theoretischer Hintergrund

In der Soziologie bietet kulturelle Kompetenz als Teil des inkorporierten Kulturkapitals großes Erklärungspotenzial für die individuelle Kulturpartizipation wie auch die intergenerationale Statusreproduktion (Georg, 2005). Nach Bourdieus Theorie der Kunstwahrnehmung zeichnet Kunst eine besondere semantische Komplexität aus (Bourdieu, 1974). Nur wer in der Lage ist, sie zu entschlüsseln, findet Gefallen an ihr. Die statusspezifische Teilhabe an den Künsten ist somit Ausdruck unterschiedlicher Kompetenzbestände in der Bevölkerung. Personen, die im Elternhaus viel kulturelle Kompetenz erworben haben, begegnen Kunst in analytisch-kritischer Distanz und können tieferliegende, durch Expert*innen verbürgte Sinnschichten erschließen. Menschen mit wenig kultureller Kompetenz sind allein auf ihr Alltagswissen angewiesen. Dadurch nehmen sie Kunstwerke oberflächlich wahr, begegnen ihnen vorwiegend emotional und sind gerade von zeitgenössischer Kunst oftmals irritiert (Hohmaier, 2015).

Fragestellung

Obwohl die Rezeptionstheorie zum Kern von Bourdieus Kultursoziologie gehört, gibt es kaum quantitative Überprüfungen (vgl. aber Rössel, 2009). Dies liegt sowohl an der uneinheitlichen Operationalisierung von Rezeption als auch der fehlenden Indikatorik zur Messung des inkorporierten Kulturkapitals. Meist werden nur indirekte oder behelfsmäßige Indikatoren eingesetzt. Grundlegende Standards klassischer Gütekriterien werden selten eingehalten (Goßmann & Mätzke, 2019) oder es werden lexikalische Wissensabfragen formuliert, die nicht zwischen kultureller Bildung und Allgemeinbildung unterscheiden. Unbeachtet bleibt das Spezifische der kulturellen Kompetenz, wie Sinnlichkeit und Symbolik (Dewey, 1934; Mollenhauer, 1990). Auch interdisziplinär finden sich wenige

elaborierte Kompetenzmessungen für den ästhetischen Bereich, z. B. KOMUS (Jordan et al., 2012) und BKKB (Frick et al., 2020) für rezeptive und KOPRA-M (Hasselhorn, 2015) für produktive Kompetenzen.

Methode

Wir stellen einen Test für rezeptive Bildkompetenz vor, der sich als domänenspezifische Operationalisierung von Bourdieus Konzept des inkorporierten Kulturkapitals versteht. Entwickelt wurde er in der Panelstudie „Kulturelle Bildung und Kulturpartizipation in Deutschland“, in der er in einem Subsample (n=227) der Wohnbevölkerung Deutschlands ab 15 Jahren 2018 zum Einsatz kam. Methodisch orientieren wir uns an der Kompetenzdiagnostik der empirischen Bildungsforschung (Hartig & Klieme, 2006). Während es dort üblich ist, Kompetenzmessungen in Classroom-Settings durchzuführen, bestand eine Herausforderung darin, ein für die breite Bevölkerung verständliches und in standardisierten Face-to-Face-Interviews zeitsparend administrierbares Instrument zu entwickeln.

Ausgehend von Panofskys kunsthistorischer Werkanalyse (Panofsky, 1939), auf die sich Bourdieus Theorie bezieht, wurde ein theoretisches Modell hergeleitet und über Leitlinien zeitgenössischer Kunstdidaktik, Schulcurricula und bestehende Kompetenzmodelle weiter fundiert (z. B. Buschkühle 2003; Kirchner & Haanstra, 2015; Wagner & Schönau, 2018). Wir unterscheiden drei Kerndimensionen rezeptiver Bildkompetenz: a) kunsthistorisches Kontextwissen, b) Analysefähigkeit und Herstellungswissen sowie c) Deutung und Transfer. Die Kompetenzen wurden an einem Set von zwölf Gemälden untersucht, die verschiedene Epochen der Kunstgeschichte prototypisch abbilden. Der Test besteht aus vier Aufgaben mit 42 Frage-Items und erforderte eine durchschnittliche Bearbeitungszeit von 23 Minuten. Die Testaufgaben wurden mit Mitteln der IRT, speziell dem Raschmodell (Rasch, 1960), skaliert und validiert.

Ergebnisse und deren Bedeutung

Es zeigt sich, dass das Messinstrument trotz seines kurzen Formats und begrenzter statistischer Power anwendungsgeeignet ist. Auf Basis generalisierter Strukturgleichungsmodelle konnte die dreidimensionale Kompetenzstruktur bestätigt werden (RMSEA = 0.032 | CFI/TLI = 0.94 | AIC = 5.898 | adj. BIC = 6001). Gegenüber Alternativmodellen mit weniger Dimensionen sowie mehr Parametern erweist sich das einparametrische Raschmodell als am besten geeignet, um die Daten zu beschreiben. Die drei Raschskalen erfüllen grundlegende Modell- und Gütekriterien der Kompetenzdiagnostik, auch wenn die Reliabilitäten aufgrund der kurzen Skalen etwas gering ausfallen (wMNSQ < 1,2, T-Wert < +2 | Pearson- χ^2-Test = insignifikant | EAP/PV-Reliabilität = 0,74, 0,66, 0,61). Gegenüber der Formalbildung, die häufig als Proxy für Kulturkapital dient, lassen sich konkrete Mechanismen der Bildrezeption identifizieren. Bourdieus Annahmen zu Erwerb und Wirkung kultureller Kompetenz werden weitestgehend bestätigt.

Kann kulturelle Bildung zum Abbau von Bildungsbarrieren beitragen? Eine längsschnittliche Analyse zur Entkopplung von sozialer Herkunft und kulturellem Kapital bei Jugendlichen

Jannis Burkhard, Stefan Kühne, Kai Maaz
DIPF

Theoretischer Hintergrund

Bourdieus (2012) These der intergenerationalen Weitergabe von kulturellem Kapital wurde inzwischen auch empirisch vielfach belegt (Kallunki & Purhonen, 2017; Raudenská & Bašná, 2021) – Eltern geben ihr kulturelles Kapital also an ihre Kinder weiter („primärer Erwerb"). Jedoch erkennt Bourdieu auch die Möglichkeit an, dass kulturelles Kapital im späteren Lebensverlauf unabhängig von der eigenen Herkunftsfamilie erworben werden kann, beispielsweise im Bildungssystem („sekundärer Erwerb"). Vor dem Hintergrund, dass ein Mangel an kulturellem Kapital eine Bildungsbarriere darstellen kann (Baumert et al., 2006), eröffnet dieser sekundäre Erwerb die Möglichkeit, Bildungsbarrieren abzubauen.

Kultureller Bildung kommt dabei eine besondere Bedeutung zu. Betrachtet man kulturelles Kapital als „knowledge of cultural symbols and ability to decode cultural messages" (Broer et al., 2019, S. 19), kann kulturelle Bildung „als (angeleiteter) Entwicklungsprozess des Subjekts in jeder einzelnen symbolischen (=Kultur-)Form verstanden werden" (Fuchs, 2008a, S. 7). Kulturelles Kapital als Kompetenz im Umgang mit kulturellen Symbolen ist demnach eine Disposition, während kulturelle Bildung einen Prozess darstellt, in dem der Umgang mit kulturellen Symbolen erlernt werden kann. Dies legt die Vermutung nahe, dass Individuen kulturelles Kapital aufbauen können, indem sie sich kulturell bilden.

Fragestellung

Der Beitrag geht der Frage nach, inwiefern kulturelle Bildung den sekundären Erwerb von kulturellem Kapital ermöglicht. Konkret sollen folgende drei Hypothesen untersucht werden:

- Der Besuch von Kursen kultureller Bildung außerhalb der Schule wirkt sich bei Jugendlichen positiv auf die Entwicklung von kulturellem Kapital über die Zeit aus.

- Der Besuch einer Schule mit musischem Schwerpunkt wirkt sich bei Jugendlichen positiv auf die Entwicklung von kulturellem Kapital über die Zeit aus.

- Diskussionen mit den Eltern über kulturelle Themen wirken sich bei Jugendlichen positiv auf die Entwicklung von kulturellem Kapital über die Zeit aus.

Methoden

Die skizzierten Fragen werden mit Daten der Startkohorte 3 des Nationalen Bildungspanels (NEPS, NEPS-Netzwerk, 2021) untersucht. Hier wurden sowohl in Welle 1 (Klasse 5, Herbst/Winter 2010) als auch in Welle 5 (Klasse 9, Herbst/Winter 2014) Konstrukte erhoben, die zur Operationalisierung von kulturellem Kapital geeignet sind (Skala „Partizipation an Hochkultur"). Somit sind längsschnittliche Analysen möglich. Insgesamt fließen Daten von n=5.778 Schüler*innen in die Analysen ein. Mit Methoden des Entropy Balancing (Hainmueller, 2012) wird der Einfluss einer Vielzahl von Kovariaten kontrolliert, sodass die Effekte der Treatment-Variablen (Besuch von Kursen kultureller Bildung außerhalb der Schule, Besuch einer Schule mit musischem Schwerpunkt, Diskussion kultureller Themen mit den Eltern) gut isoliert werden können. Um eine bestmögliche und repräsentative Parameterschätzung zu gewährleisten, wurden fehlende Werte multipel imputiert und Survey-Gewichte verwendet. Alle Schritte zur Datenaufbereitung und -analyse wurden in der Statistik-

Software R (R Core Team, 2022) mit den Paketen tidyverse (Wickham et al., 2019), mice (van Buuren & Groothuis-Oudshoorn, 2011) und MatchThem (Pishgar et al., 2021) vorgenommen.

Ergebnisse und deren Bedeutung

Vorläufige Ergebnisse legen nahe, dass der Besuch von Kursen kultureller Bildung die Entwicklung des kulturellen Kapitals signifikant positiv beeinflusst (ATT 2,60). Für den Besuch von Schulen mit musischem Schwerpunkt (ATT 0,53) sowie Diskussionen mit den Eltern (ATT 0,51) fallen die Effekte deutlich kleiner aus und sind nicht statistisch signifikant. Kursbesuche ermöglichen es den Jugendlichen also, kulturelles Kapital aufzubauen, während dies für die anderen beiden Variablen nicht nachgewiesen werden kann. Jugendliche, für die ein Mangel an kulturellem Kapital eine Bildungsbarriere darstellt, können diese Bildungsbarriere daher vermutlich zumindest teilweise abbauen, indem sie Kurse kultureller Bildung besuchen. Vor dem Hintergrund der These, dass kulturelles Kapital auch soziale Mobilität ermöglichen kann (DiMaggio, 1986), ist dieses Ergebnis auch unter der Perspektive von Bildungsgerechtigkeit relevant. Im Rahmen der GEBF werden die finalen Ergebnisse präsentiert. Außerdem wird die Operationalisierung der beiden zentralen Konstrukte im Kontext der anderen Beiträge des Symposiums kritisch diskutiert.

B-12: Determinanten des subjektivem Wohlbefindens im Kinder- und Jugendalter

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* T03 R02 D26

Determinanten des subjektivem Wohlbefindens im Kinder- und Jugendalter

Chair(s): **Ricarda Steinmayr** (Technische Universität Dortmund), **Maike Luhmann** (Ruhr Universität Bochum)
*Diskutant*in(nen):* **Fani Lauermann** (Technische Universität Dortmund)

Nach Diener et al. (2018) wird das SWB durch eine allgemeine subjektive Lebenszufriedenheit, die Anwesenheit positiver und die Abwesenheit negativer Emotionen sowie der Zufriedenheit in verschiedenen Lebensbereichen definiert. In Studien mit Kindern und Jugendlichen konnten bedeutsame Zusammenhänge zwischen dem SWB und verschiedenen Indikatoren der psychischen und physischen Gesundheit (z.B. Jonsson et al., 2017) sowie Leistungskriterien gezeigt werden (Steinmayr et al., 2018). Während vor allem querschnittliche Untersuchungen die Bedeutung des SWB auch im Kindes- und Jugendalter darlegen, sind jedoch viele Fragen in diesem Altersbereich weniger beforscht, z.B. was Veränderungen im SWB determiniert (s. aber Steinmayr et al., 2019). Aus Studien an Erwachsenen sind einige Faktoren bekannt, die interindividuelle Veränderungen in intraindividuellen Verläufen im SWB erklären können, z.B. Persönlichkeit, Selbstwirksamkeit, etc. (Luhmann, 2017). Um diese Determinanten zu identifizieren, benötigt man jedoch längsschnittliche Studien, die es in im Kinder- und Jugendbereich kaum gibt. Darüber hinaus weiß man, dass der „person-environment-fit“ Hypothese zufolge u.a. das SWB von Personen durch eine Passung von individuellen Eigenschaften und Eigenschaften der Umwelt determiniert wird (Hunt, 1975). Unsere Umwelt ändert sich jedoch ständig, vor allem vor dem Hintergrund der aktuellen Großereignisse wie Pandemie, Krieg in Europa, Klimakrise etc.. Somit ist bei der Untersuchung von Determinanten des SWBs nicht nur ein längsschnittlicher Ansatz zu verfolgen, sondern auch zeitliche Veränderungen sowie Unterschiede zwischen Kohorten zu beachten. Im vorliegenden Symposium wird zunächst eine qualitative Studie vorgestellt, in der Jugendliche im Sommer 2022 dazu befragt wurden, was sie glücklich macht. Anschließend wird eine längsschnittliche Studie präsentiert, die sich mit den Konsequenzen romantischer Beziehungen im Jugendalter für die Entwicklung des Selbstwertgefühls, einer der wichtigsten Determinanten des SWBs, und in einem geplantem Kohorten-Sequenz-Design mit der Frage beschäftigt, ob sich das Wohlbefinden jugendlicher Singles über die letzten 10 Jahre verändert hat. Im dritten Beitrag wird vor dem Hintergrund des wahrgenommenen gestiegenen Leistungsdrucks bei Schüler*innen der Frage des Zusammenhangs von Leistungsdruck und SWB bei Grundschulkindern der 4. Klasse nachgegangen. In der letzten Studie wird die Entwicklung des SWBs von Grundschulkindern untersucht. Dabei werden sowohl das Geschlecht, Migrationshintergrund, SES sowie ihre Intersektionen als auch das Sozialklima als Moderatoren der individuellen Entwicklungsverläufe im SWB betrachtet. Darüber hinaus werden Ausgangswert und Entwicklungsfaktor des SWBs in Bezug auf ihren Zusammenhang mit schulischen Kompetenzen untersucht. Das Symposium bietet somit bedeutsame Einblicke in Determinanten des SWBs und seiner Entwicklung im Kindes- und Jugendalter aus persönlichkeitspsychologischer, pädagogisch-psychologischer und erziehungswissenschaftlicher Perspektive. Die Beiträge werden anschließend von Prof.‘in Dr. Fani Lauermann diskutiert.

Beiträge des Symposiums

Subjektive Annahmen zu den Bedingungen von Wohlbefinden bei Jugendlichen am Anfang der Pubertät

Alicia Neumann, Marcus Roth
Universität Duisburg-Essen

Theoretischer Hintergrund

Obwohl ein klarer Zusammenhang zwischen subjektivem Wohlbefinden (SWB) und einer Vielzahl unterschiedlicher Indikatoren der psychischen und physischen Gesundheit nachgewiesen wurde (Überblick bei Diener et al., 2018), existieren bislang nur wenige Studien, die Determinanten des Wohlbefindens speziell bei Jugendlichen untersuchen. Bisherige Studien legen nahe, dass es von der frühen bis zur mittleren Adoleszenz eine durchschnittliche Abnahme des subjektiven Wohlbefindens gibt. Shek und Liu (2014) konnten in einer Hongkonger-Studie zeigten, dass die Lebenszufriedenheit von Jugendlichen im Alter von 12,5 bis 14,5 Jahren durchschnittlich abnahm. Diese Abnahme war ausgeprägter für Jungen und für Jugendliche aus nicht intakten Familien. González-Carrasco et al. (2017) fanden in einer Stichprobe von spanischen 10- bis 15-jährigen Jugendlichen ebenfalls eine Abnahme des SWB ab dem Alter von 11 bis 12 Jahren, hier aber war die Abnahme bei Mädchen ausgeprägter als bei Jungen. In Deutschland sind keine solchen längsschnittlichen Studien bekannt, die diese Abnahme belegen, jedoch gibt es querschnittliche Studien, welche Jugendliche unterschiedlichen Alters miteinander vergleichen und die längsschnittlichen Befunde stützen (z.B. Goldbeck et al., 2017). Was jedoch fehlt, sind Befunde zu den Faktoren, welche unterschiedliche Verläufe des subjektiven Wohlbefindens bei Jugendlichen vorhersagen. Es ist aber von besonderer praktischer Relevanz, Determinanten für eine optimale Entwicklung des Wohlbefindens in diesem Alter zu identifizieren, denn gerade die Pubertät ist gekennzeichnet von körperlichen, kognitiven, sozialen und emotionalen Veränderungen, die mit einer hohen Anpassungsleistung einhergehen (Berk, 2020).

Fragestellung

Um Risikofaktoren identifizieren zu können, die die Abnahme des SWB in der Adoleszenz begünstigen oder vorantreiben, wurde in einem ersten Schritt ein qualitativer Ansatz gewählt, in dem zunächst die subjektive Sicht der Jugendlichen selbst exploriert und nach den Annahmen gefragt wurde, welche Faktoren für das eigene Wohlbefinden ausschlaggebend sind. In Interviews mit 13-16-jährigen Jugendlichen wurden daher offene Fragen über ihre Definitionen von Glück und Wohlbefinden gestellt und erfragt, welche Bedingungen sie für persönliches Wohlbefinden als notwendig erachten (subjektive Annahmen zu den Bedingungen des Wohlbefindens).

Methode

Es wurden semi-strukturierte Interviews mit 44 Jugendlichen unterschiedlicher weiterführender Schulformen (männlich: 43%, weiblich: 57%) zwischen 13 und 15 Jahren (M = 13.73) geführt. Es gab zwei große Fragenkomplexe. Der erste beschäftigte sich mit dem Verständnis von subjektiven Wohlbefinden und wurde mit der Globalfrage: „Was ist das für Dich, Glück oder Wohlfühlen oder Zufriedenheit?“ erfragt. Der zweite Fragenkomplex beschäftigte sich mit den Bedingungen und Determinanten um glücklich zu sein und wurde mit der Globalfrage: „Was sind Faktoren und Bedingungen/„Must haves“ zum glücklich Sein?“ abgefragt. Anschließend wurden die Interviews

transkribiert und in Anlehnung an die Inhaltsanalyse nach Mayring (2022) codiert. Aus den ursprünglichen zwei Globalfragen wurde aufgrund von zu großer Überschneidung der Antworten eine übergeordnete Frage („Was brauche ich, um glücklich zu sein?“) erstellt.

Ergebnisse

Die Antworten der Jugendlichen zeigten eine große Varianz und waren recht detailliert, was den richtigen Einsatz eines qualitativen Ansatzes unterstreicht. Insgesamt konnten 54 distinkte Kategorien identifiziert werden. Die am häufigsten genannten Kategorie war dabei die der positiven Emotionen („Jemand der glücklich ist, natürlich, und der ein Lächeln im Gesicht hat und der höflich ist und der freundlich aussieht.“); die am wenigsten genannte Kategorie war Dankbarkeit. Zudem wurden häufig Faktoren wie soziale Beziehungen, materieller Wohlstand, Persönlichkeitsaspekte, Erfolg, Freiheitsgefühle und eine hedonistische Ansicht von Glück genannt. Überraschenderweise wurde Spiritualität/Religiosität von keinem Jugendlichen genannt, obwohl dieser Faktor in der Glücksforschung bisher eine große Bedeutung gespielt hat.

Diskussion und Relevanz

Die Ergebnisse ließen Rückschlüsse auf Geschlechtsunterschiede sowie Zuwanderungsgeschichte zu. Darüber hinaus geben sie Hinweise, dass bestimmte Korrelate, z.B. Religiosität, des subjektiven Wohlbefindens und dessen Entwicklung aufgrund gesellschaftlicher Veränderungen ihre Bedeutung verloren haben. Dies muss in zukünftigen Studien bzw. Re-Analysen von bestehenden Panel-Untersuchungen quantitativ-empirisch untersucht werden.

Singleleben in der Adoleszenz – Selbstwertentwicklung und historische Veränderung im Wohlbefinden

Tita Gonzalez Avilés
Friedrich-Schiller-Universität Jena

Theoretischer Hintergrund

In der Adoleszenz machen viele Jugendliche ihre ersten romantischen Erfahrungen und gehen Partnerschaften ein. So geben im Alter von 16 Jahren etwa die Hälfte der Jugendlichen (48.4 – 52.3%) an, schon einmal eine romantische Beziehung gehabt zu haben (Boisvert & Poulin, 2016; Connolly et al., 2013). Jedoch bleiben rund 22% über das Jugendalter hinweg Single und gehen ihre erste Beziehung zu einem späteren Zeitpunkt ein (Gonzalez Avilés et al., 2021). Aufgrund der normativen und bedeutsamen Rolle romantischer Beziehungen in dieser Altersspanne, könnte sich das Ausbleiben der ersten romantischen Beziehung negativ auf den Selbstwert auswirken.

Gleichzeitig ist das Singleleben in den vergangenen 50 Jahren akzeptierter und normativer geworden: der Anteil Singles ist in westlichen Kulturen gestiegen (Pepping et al., 2018) und gesellschaftliche Normen bzgl. romantischer Beziehungen haben sich verändert (Bühler & Nikitin, 2020). Die Ehe wird seltener angestrebt, Bedürfnisse nach Individualismus und Autonomie haben zugenommen und später geborene Kohorten sind generell skeptischer gegenüber sozial etablierten Strukturen. Diese historischen Veränderungen könnten sich auch im Wohlbefinden junger Singles zeigen, so dass später geborene Singles zufriedener mit ihrem Singleleben und ihrem Leben im Allgemeinen sind, als früher geborene Singles.

Fragestellungen

In der Studie sollen zwei zentrale Fragestellungen rund ums Singleleben in der Adoleszenz untersucht werden: (1) Wie wirkt sich das Ausbleiben einer romantischen Beziehung im Jugendalter auf den Selbstwert aus? (2) Hat sich das Wohlbefinden jugendlicher Singles über die letzten 10 Jahre verändert?

Methode

Für beide Fragestellungen werden Daten aus einer repräsentativen Längsschnittstudie pairfam aus Deutschland herangezogen, in der seit 2008 jährlich der Selbstwert (3 Items), die Zufriedenheit mit dem Singleleben (1 Item) sowie die Lebenszufriedenheit (1 Item) erfasst wurden. Insgesamt nahmen N = 4,338 Jugendliche (M = 16.22) bei der ersten Erhebung in den Jahren 2008/2009 teil. Zusätzlich wurde in den Jahren 2018/2019 eine weitere Stichprobe von N = 2,476 Jugendlichen rekrutiert.

Für die Beantwortung der ersten Fragestellung wurden Daten von N = 1,395 Jugendlichen über 10 Messzeitpunkte hinweg analysiert, die zum Zeitpunkt des ersten Messzeitpunkts noch keine erste Beziehung hatten. Anhand piecewise growth curve models (Duncan et al., 2013) wurde die Selbstwertentwicklung sowohl vor als nach der ersten Beziehung modelliert. Der moderierende Effekt vom Alter der ersten Beziehung wurde mittels Interaktionsterm getestet.

Zur Beantwortung der zweiten Fragestellung sollen Daten von N = 1,927 Jugendlichen analysiert werden, die über 3 Messzeitpunkte hinweg Single waren. Anhand eines Kohorten-Sequenz-Designs soll untersucht werden ob sich früher geborene (2008/2009: Geburtsjahre 1991-1993) von später geborenen jugendlichen Singles (2018/2019: Geburtsjahre 2001-2003) in ihrem Wohlbefinden (Zufriedenheit mit dem Singleleben & Lebenszufriedenheit) unterscheiden. Veränderungen im Wohlbefinden über die Singlezeit sollen ebenfalls modelliert werden.

Ergebnisse

Ergebnisse der ersten Fragestellung zeigen, dass sich Personen, die den Beginn ihrer ersten Beziehung bis ins junge Erwachsenenalter (M = 26.22) verzögern, in ihrer Selbstwertentwicklung nicht von Personen unterscheiden, die bereits früher ihre erste Beziehung eingegangen sind: Der Selbstwert blieb unabhängig vom Alter der ersten Beziehung stabil in den Jahren vor der ersten romantischen Beziehung. Somit scheint das Singleleben im Jugendalter kein Risiko für die Selbstwertentwicklung darzustellen. Möglicherweise liegt das Ausbleiben einer romantischen Beziehung bis Mitte 20 noch im normativen Rahmen. Nachteile für den Selbstwert könnten erst bei weiterer Verzögerung der ersten Beziehung ersichtlich werden.

Zum Zeitpunkt der Einreichung ist eine Präregistrierung für die zweite Fragestellung erstellt und die dazugehörigen Daten wurden aufbereitet.

Relevanz

Die Befunde beider Fragestellungen sollen zu einem tieferen Verständnis vom Singleleben im Jugendalter führen und deren Bedeutung für die Persönlichkeitsentwicklung und das Wohlbefinden unterstreichen.

Vergleich des elterlichen und schulischen Leistungsdrucks hinsichtlich ihres Zusammenhangs mit dem schulischen Wohlbefinden von Schülerinnen und Schülern

Ruben Kleinkorres, Justine Stang-Rabrig, Nele McElvany
Technische Universität Dortmund

Theoretischer Hintergrund

Das Wohlbefinden von Schülerinnen und Schülern (SuS) gilt als wichtiges Kriterium des Schulerfolgs (Kanonire et al., 2020; OECD, 2017). Das multidimensionale Konstrukt konstituiert sich aus psychischen, physischen und sozialen Facetten (WHO, 2014). Durch die Bedeutung, die dem schulischen Wohlbefinden beigemessen wird, ist es wichtig, neben Faktoren, die es positiv beeinflussen auch solche zu identifizieren, die negativ damit zusammenhängen. In den letzten zwei Jahrzehnten hat der Anteil der SuS, die über hohen Leistungsdruck berichten, in vielen europäischen Ländern zugenommen (Löfstedt et al., 2020). Externer Leistungsdruck ist mit niedrigerem Wohlbefinden bei adoleszenten Schulkindern assoziiert (Deb et al., 2015; Torsheim & Wold, 2001), was durch einen Mangel an Autonomie und Selbstbestimmung erklärt werden könnte (Ryan & Deci, 2020). Schulbezogener Leistungsdruck wird häufig durch Eltern oder die Schule ausgeübt und zeichnet sich beispielsweise durch hohe Leistungserwartungen oder strenge Anforderungen bezüglich der Erledigung von Schularbeiten aus (Hirvonen et al., 2019; Quach et al., 2015). Da die Eltern die Hauptakteure der frühkindlichen Sozialisation sind (Grusec & Hastings, 2007), ist es wahrscheinlich, dass sie während der Grundschulzeit noch einen stärkeren Einfluss haben als das schulische Umfeld, sodass der von ihnen ausgeübte Druck für das Wohlbefinden der Kinder bedrohlicher ist. Der Zusammenhang zwischen Leistungsdruck und Wohlbefinden wurde jedoch meist in adoleszenten Schülerpopulationen untersucht (Deb et al., 2015; Torsheim & Wold, 2001) und ein Vergleich verschiedener Quellen des Leistungsdrucks hinsichtlich ihres Zusammenhangs mit verschiedenen Wohlbefindensfacetten wurde bisher nicht durchgeführt. Der vorliegende Beitrag widmet sich daher der Untersuchung der Zusammenhänge zwischen elterlichem beziehungsweise schulischem Leistungsdruck und zentralen Wohlbefindensfacetten von SuS in der Grundschule als bedeutsamer Schulphase.

Fragestellungen

In der ersten Forschungsfrage (F) der vorliegenden Studie wurde untersucht, wie elterlicher und schulischer Leistungsdruck jeweils mit verschiedenen Wohlbefindensfacetten (Schulzufriedenheit, Schulfreude, Sorgen wegen der Schule, körperliche Beschwerden, soziale Eingebundenheit) zusammenhängen (F1a) und ob die Zusammenhänge stabil bleiben, wenn für individuelle Variablen (Geschlecht, akademische Leistung) kontrolliert wird (F1b). Im Anschluss daran wurde überprüft, ob der elterliche Leistungsdruck stärker mit den einzelnen Wohlbefindensfacetten zusammenhängt als der schulische Leistungsdruck (F2a). Auch hier wurde untersucht, ob das Befundmuster robust gegenüber dem Einbezug der Kontrollvariablen ist (F2b).

Methode

Die Analysen basieren auf Feldtestdaten der internationalen Grundschul-Lese-Untersuchung (IGLU). Die Stichprobe umfasste N = 342 Viertklässlerinnen und Viertklässler (45.9% weiblich, MAlter = 10.29 Jahre [SD = 0.54]). Die Teilnehmenden machten dabei Angaben zu ihrer Schulzufriedenheit (1 Item),

Schulfreude (3 Items; Cronbachs □: .75), Sorgen wegen der Schule (4 Items; Cronbachs □: .82), körperlichen Beschwerden wegen der Schule (4 Items; Cronbachs □: .83) und ihrer sozialen Eingebundenheit im Schulkontext (3 Items; Cronbachs □: .80). Die Datenaufbereitung, Berechnung der Deskriptiva und Durchführung der Analysen erfolgten in R (Version 3.6.3; R Core Team, 2020).

Ergebnisse

Latente Strukturgleichungsmodelle ergaben, dass sowohl elterlicher (□ = .34) als auch schulischer Leistungsdruck (□ = .33) signifikant positiv mit Sorgen wegen der Schule zusammenhingen. Diese Zusammenhänge waren zudem robust gegenüber der Aufnahme der Kontrollvariablen. Elterlicher Leistungsdruck war darüber hinaus signifikant negativ mit Schulzufriedenheit (□ = -.26) und sozialer Eingebundenheit (□ = -.25) assoziiert, während sich ein signifikant positiver Zusammenhang mit physischen Problemen wegen der Schule (□ = .26) zeigte. Von diesen Assoziationen war allerdings nur jene mit der Schulzufriedenheit robust gegenüber der Aufnahme der Kontrollvariablen (□ = -.17). Bei der gleichzeitigen Betrachtung von elterlichem und schulischem Leistungsdruck zeigte sich ein stärkerer Zusammenhang der Schulzufriedenheit mit dem elterlichen Leistungsdruck. Nach Aufnahme der Kontrollvariablen waren die Zusammenhänge jedoch nicht mehr signifikant und blieben nur in ihrer Tendenz bestehen.

Relevanz

Die vorliegende Studie erweitert durch simultane Fokussierung mehrerer Wohlbefindensfacetten sowie zweier Quellen von Leistungsdruck systematisch den bisherigen Kenntnisstand. Das Befundmuster kann beispielsweise sozialisationstheoretisch erklärt werden (z.B. Grusec & Hastings, 2007). Implikationen für Forschung und Praxis werden diskutiert.

School should be fun (but is not fun for everyone). Die Bedeutung von Geschlecht und Migrationshintergrund und ihrer Intersektion für die Entwicklung des Wohlbefindens in der Grundschule

Rebecca Lazarides[1], Elisa Oppermann[2], Charlott Rubach[3], Sandi Simpkins[4]
[1]Universität Potsdam, [2]Universität Bamberg, [3]Universität Rostock, [4]University of Irvine, LA, USA

Theoretischer Hintergrund

Schulisches Wohlbefinden kann als Zielkriterium schulischer Bildungsprozesse betrachtet werden (Clement, 2010). Unter Wohlbefinden werden dabei positive Emotionen und Kognitionen und die Abwesenheit negativer Emotionen und Kognitionen verstanden (Hascher, 2004). Für das Jugendalter ist gut belegt, dass Wohlbefinden von Bedeutung für akademische Leistungen ist (Hascher & Hagenauer, 2011). Dabei ist bekannt, dass Jungen ein besonders geringes Wohlbefinden im schulischen Kontext berichten (Palsdottir et al., 2012), allerdings ist wenig bekannt darüber inwiefern das Geschlecht Lernender mit anderen sozialen Kategorien wie dem Migrationshintergrund oder der sozio-ökonomischen Herkunft (SES) interagiert. Zudem bestätigen metaanalytische Befunde Zusammenhänge zwischen Wohlbefinden und Klassenklima (Wang et al., 2020). Allerdings beziehen sich viele Studien auf das Jugendalter und wenig ist über die Entwicklung von schulischem Wohlbefinden bei Grundschüler:innen bekannt.

Fragestellungen

Die vorliegende Studie geht der Frage nach, wie sich das Wohlbefinden von Grundschulkindern im Verlauf der dritten und vierten Jahrgangsstufe entwickelt und welche Zusammenhänge sich zwischen den individuellen Merkmalen Geschlecht, Migrationshintergrund und SES sowie ihren Intersektionen in Bezug auf den Ausgangswert und die Entwicklung des Wohlbefindens der Grundschüler:innen zeigen. Zudem wird untersucht, welche Bedeutung das von den Lernenden wahrgenommene Sozialklima für den Ausgangswert und die Entwicklung des Wohlbefindens haben. Außerdem gehen wir der Frage nach, inwiefern der Ausgangswert und die Entwicklung des Wohlbefindens in der Grundschule für die Kompetenzen in den Bereichen Mathematik und Lesen am Ende der vierten Jahrgangsstufe von Bedeutung sind.

Methode

Genutzt wurden Daten von N = 2,162 Grundschüler:innen (MAlterT1 = 9.28 Jahre, SD = 0.45; 47.8% Mädchen) die an der BIKS Studie (Homuth et al., 2014) zu mindestens einem von drei Messzeitpunkten teilnahmen. Die Mehrheit der Teilnehmenden gaben die Sprache Deutsch als Familiensprache an (82.3%). Die meisten Lernenden berichteten, dass beide Eltern in Deutschland geboren wurden (76.8%). Der SES der Familie wurde mit dem Highest International Socio-Economic Index of Occupational Status (HISEI; Ganzeboom et al., 1992) gemessen und war vergleichbar mit der repräsentativen Stichprobe der PISA 2018 Studie (BIKS Teilsample: M = 50.32, SD = 16.42; PISA 2018: M = 50.34, SD = 16.36) (Reiss et al., 2019).

Ergebnisse

Lineare Wachstumskurvenmodelle zeigten ein signifikantes Absinken des Wohlbefindens von der dritten zur vierten Jahrgangsstufe. Jungen und Kinder, die Deutsch als Familiensprache angaben, berichteten ein besonders geringes Wohlbefinden zu Beginn der dritten Jahrgangsstufe. Jungen, die angaben, dass beide Elternteile in Deutschland geboren wurden, berichteten ein besonders geringes Wohlbefinden im Gruppenvergleich und Mädchen, die angaben, dass mindestens ein Elternteil im Ausland geboren wurde, berichteten ein besonders hohes Wohlbefinden zu Beginn der dritten Jahrgangsstufe. Das von den Lernenden berichtete Sozialklima in der Schulklasse zu Beginn der dritten Jahrgangsstufe korrelierte positiv mit dem Ausgangsniveau des Wohlbefindens der Kinder zu Beginn der dritten Jahrgangsstufe. Die Entwicklung des Wohlbefindens von der dritten zur vierten Jahrgangsstufe stand in positivem Zusammenhang mit der Kompetenz in Mathematik und mit der Kompetenz im Bereich Lesen in der Mitte der vierten Jahrgangsstufe. Der Ausgangswert des Wohlbefindens zu Beginn der dritten Jahrgangsstufe stand in positivem Zusammenhang zur Mathematikkompetenz in der Mitte der vierten Jahrgangsstufe. Die Ergebnisse verdeutlichen die Bedeutung der Entwicklung des Wohlbefindens für Kompetenzbeginn bereits im Grundschulalter. Zudem wird die Wichtigkeit der Berücksichtigung multipler sozialer Kategorien und ihrer Interaktionen deutlich.

Relevanz

Die Zusammenhänge zwischen sozialen Kategorien und der Entwicklung des individuellen Wohlbefindens wurde bislang selten im Grundschulalter untersucht. Besonders vor dem Hintergrund der gesellschaftlichen Herausforderung der COVID-19 Pandemie ist jedoch deutlich geworden, dass

Wohlbefinden in der Schule ein zentrales Kriterium für Bildungserfolg darstellt. Mangelndes Wohlbefinden kann zu Schulabsenz und sinkendem Kompetenzgewinn beitragen. Die vorliegende Studie trägt dazu bei, die problematischen Konsequenzen geringen Wohlbefindens bei Grundschüler:innen besser zu verstehen und das Zusammenwirken individueller Merkmale in Bezug auf die Ausgangslage des Wohlbefindens im Grundschulalter besser nachzuvollziehen.

B-13: Herausforderungen und Entwicklungen bei der Messung schriftsprachlicher Fähigkeiten

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* T03 R02 D39

Herausforderungen und Entwicklungen bei der Messung schriftsprachlicher Fähigkeiten

Chair(s): **Johannes Hartig** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany)
*Diskutant*in(nen):* **Astrid Neumann** (Leuphana-Universität Lüneburg)

Schriftsprachliche Fähigkeiten sind als eine Grundvoraussetzung für Lernen unabdingbar für Bildungserfolg, aber darüber hinaus auch zentral für Teilhabe in vielen gesellschaftlichen Bereichen. Ihre Messung ist folgerichtig ein intensiv beforschter Gegenstand der empirischen Bildungsforschung, in dem verschiedene fachliche Perspektiven fruchtbar zusammenarbeiten. Aufgrund der Komplexität des Gegenstandes, der u.a. durch die Vielfalt sprachlicher Anwendungskontexte, verschiedene Textsorten, die Mehrdimensionalität rezeptiver und produktiver Fähigkeiten und nicht zuletzt die Beschäftigung mit verschiedenen Mutter-, Fremd- und Unterrichtssprachen bedingt ist, bleibt das Forschungsfeld trotz der langjährigen Beforschung dynamisch und aktiv. Im Symposium werden Innovationen und Entwicklungen bei der Messung schriftsprachlicher Fähigkeiten aus verschiedenen fachlichen Perspektiven (Sprachdidaktik, Erziehungswissenschaft, Psychologie) vorgestellt. Die Beiträge beschäftigen sich sowohl mit produktiven als auch rezeptiven Fähigkeiten und haben Englisch als Fremdsprache, sprachübergreifende Schreibkompetenzen, und Deutsch als Unterrichtssprache im Blick.

Der erste Beitrag befasst sich mit der Diagnose argumentativer Kompetenzen in Englisch. Differenzierte Ratings eines großen Korpus von Schülertexten werden verwendet, um die Messung der Kompetenzen der Schüler*innen zu verbessern und die Diagnosekompetenzen von Englischlehrpersonen zu trainieren. Im zweiten Beitrag wird die Frage untersucht, inwieweit Fähigkeiten mehrsprachiger Schüler*innen durch ein integriertes Modell mehrsprachigen Schreibens beschrieben werden können, das eine sprachübergreifende Schreibkompetenzdimension beinhaltet. Untersucht werden hierzu Daten mehrsprachiger Schüler*innen mit russisch-deutschem und türkisch-deutschem Sprachhintergrund. Im dritten Beitrag wird die Bearbeitung von integrierten Aufgaben (reading into writing) in Englisch als Fremdsprache untersucht, er ist somit an der Schnittstelle zwischen Lesen und Schreiben angesiedelt. Auf Basis von bei computerbasierten Aufgabenbearbeitungen aufgezeichneten Verhaltensdaten wird untersucht, wie Annotationsstrategien (Hervorhebung und Notizen) eingesetzt werden und wie sie mit der Schreibleistung von Schüler*innen und Studierenden zusammenhängen. Der vierte Beitrag stellt die Entwicklung eines Bildungsstandard-basierten Lesetests vor, bei der die Mehrdimensionalität sowohl der individuellen Fähigkeiten als auch der sprachlichen Anwendungskontexte berücksichtigt werden soll. Der Test wird in Österreich sowohl für Unterrichts- und Schulentwicklung als auch zum Bildungsmonitoring verwendet, womit in diesem Beitrag auch die Anwendung sprachlicher Diagnostik in der Bildungspraxis behandelt wird.

Die Ergebnisse der in den Beiträgen vorgestellten Forschung können zur Weiterentwicklung der Erfassung schriftsprachlicher Fähigkeiten beitragen. Ein abschließender Diskussionsbeitrag adressiert, in welcher Hinsicht dies gelingt, welche Limitationen die vorgestellten Arbeiten haben und wo noch weiterer Forschungsbedarf in diesem Themengebiet angedacht werden kann.

Beiträge des Symposiums

Differenzierte Diagnose englischer Schreibkompetenzen auf der Oberstufe

Stefan Keller[1], Julian Lohmann[2], Jens Möller[2]
[1]Pädagogische Hochschule Zürich, [2]Universität zu Kiel

Theoretischer Hintergrund

Schreiben in der Fremdsprache Englisch gehört zu den zentralen Bildungszielen der gymnasialen Oberstufe und ist relevant für Studierfähigkeit und employability in einer zunehmend globalisierten Welt (Autor, 2013). Insbesondere das argumentative Schreiben nimmt dabei einen hohen Stellenwert ein. Es kann als Fähigkeit verstanden werden, die eigene Meinung zu einem Thema adressatenspezifisch zu strukturieren und mittels grammatischer und lexikalischer Mittel differenziert darzustellen, sodass Leserinnen und Leser wichtige Punkte erkennen und sich merken können (CEFR, 2001).

In den letzten Jahren sind in europäischen Ländern erste empirische Studien argumentativen Schreiben in Englisch entstanden (Schoonen et al., 2010). Diese untersuchten die Schreibfähigkeit in Englisch mit einem holistischen Urteil, z.B. in Abhängigkeit von Moderatorvariablen wie Geschlecht, sprachlicher und sozialer Hintergrund usw. untersuchten (Author, 2020). Eine besondere Herausforderung ist dabei die differenzierte und reliable Erfassung von Textqualität jenseits eines holistischen Gesamturteils. Es existieren zwar differenzierte Modelle von argumentativer Schreibkompetenz (Hyland, 1990), wobei insbesondere die Bereiche „Sprachqualität“ (z.B. Grammatik, Lexis, Rechtschreibung), Struktur (z.B. Absatzgestaltung, Argumentstrukturierung) und Inhalt (z.B. Klarheit und Relevanz der themenspezifischen Aussagen) bedeutsam sind. Bislang ist es jedoch nicht gelungen, differenzierte und reliable Beurteilungen solcher genre-spezifischen Texteigenschaften so zu modellieren, welche für large-scale Analysen nutzbar sind.

Fragestellung

Wie lassen sich differenzierte, genre-spezifische Qualitätsmerkmale bei Texten des argumentativen Schreibens in der Fremdsprache Englisch erfassen? Lassen sich insbesondere die Bereiche „Sprachqualität“, „Struktur“ und „Inhalt“ empirisch trennen und reliabel modellieren?

Methode

Für die differenzierte Analyse des argumentative writing wurde ein Analyseraster mit 25 Kategorien erstellt. 10 Items betrafen den Textaspekt „Sprache“, 5 Items betrafen „Struktur“ und 10 Items betrafen „Inhalt“. Als Grundlage für diese Analysekriterien diente ein bestehendes Instrument zur Analyse von Texten aus der Studieneingangsstufe in der Schulsprache, welche für englisches Schreiben angepasst wurde (Bremerich-Vos & Scholten-Akoun, 2016). Mit diesem Raster wurden N = 2289 Texte von Schülerinnen und Schülern aus dem 11. Schuljahr zu zwei Schreibprompts aus dem TOEFL Test analysiert („independent writing“; Rupp et al., 2019). Das Rating erfolgte durch Raterinnen und Rater, welche für die Anwendung des Analyserasters über einen längeren Zeitraum geschult wurden. Dabei wurden teilweise auch Schärfungen und Überarbeitungen an den einzelnen Items vorgenommen. Nach der Schulungsphase wurde jeder Text von zwei Raterinnen und Ratern mit dem finalen Beurteilungsinstrument beurteilt und die Inter-Rater-Reliabilität geprüft.

Ergebnisse und ihre Bedeutung

Die Ergebnisse des Ratings zeigten, dass sich für viele der entwickelten Kategorien reliable Urteile herstellen liessen (d.h. Kappa ≥ .5). Bei einigen binären Kodierungen waren die Werte stark asymmetrisch verteilt und liessen keine sinnvollen Rückschlüsse über die Textqualität zu (z.B. Variabilität des Wortschatzes). Bei anderen Dimensionen liess sich die Inter-Rater-Reliabilität durch Ausschluss eines 'schlechten' Raters verbessern. Insgesamt gelang es für die Oberkategorien Sprachqualität (.747), Struktur (.723) und Inhalt (.641) Urteile mit zufriedenstellender Inter-Rater-Reliabilität zu finden.

Durch die detaillierte Analyse von Schreibkompetenzen werden erstens differenzierte Analysen der Schreibkompetenzen von Jugendlichen ermöglicht, welche auch fachdidaktische Relevanz haben. Beispielsweise können Profile von Schreibkompetenzen gebildet oder überprüft werden, wie sich diese Faktoren im Unterricht entwickeln. Zweitens lassen sich die Ratings auch als Benchmark-Scores für Studien zur Beurteilungsqualität von Lehrpersonen nutzen. Dabei werden aus dem Gesamtkorpus randomisiert Texte gezogen und dann Probandinnen und Probanden zur Beurteilung vorgelegt, wobei teacher judgment accuracy (Coladarci, 1986) untersucht werden kann. Drittens besteht die Möglichkeit zur Automatisierung dieser Ratings mittels Algorithmen des natural language processing (Deane, 2013), womit in Zukunft auch automatisierte Kompetenzmessung oder Feedback möglich werden. Diese Perspektiven werden in dem Beitrag kritisch diskutiert.

Modellierung von mehrsprachiger Schreibfähigkeit als integriertes Konstrukt

Birger Schnoor, Irina Usanova
Universität Hamburg

Theoretischer Hintergrund

Internationale Schülerleistungsvergleichsstudien zeigen regelmäßig eine Bildungsbenachteiligung von Schülerinnen und Schülern mit Migrationshintergrund gegenüber der einheimischen Referenzbevölkerung (z. B. OECD, 2019). Ursächlich hierfür ist vor allem die geringere Ressourcenausstattung ihrer Familien, entstanden durch die disruptive Wirkung von Migration auf ökonomisches, kulturelles und soziales Kapital (Nauck & Schnoor, 2015). Migration birgt aber auch Potenziale, wobei eines die migrationsbedingte Mehrsprachigkeit ist, die als Humankapital im Arbeitsmarkt dienen kann (Agirdag, 2014a, 2014b; Gándara, 2018; Rumbaut,2014).

Die Erforschung von Bildungsvorteilen migrationsbedingter Mehrsprachigkeit ist ein wachsendes Forschungsfeld. Die groß angelegten internationalen Schülerleistungsvergleichsstudien (z. B. PISA; PIRLS) messen hauptsächlich rezeptive Fähigkeiten (Lesen) in der Majoritätssprache. Wenn sie Mehrsprachigkeit berücksichtigen, dann wird diese nicht direkt durch Leistungstests erfasst, sondern durch Ersatzvariablen (z. B. Selbsteinschätzungen oder Sprachpraktiken) erhoben. Unterrepräsentiert sind Studien zum mehrsprachigen Schreiben bei Jugendlichen anhand substanzieller Stichproben (Gogolin et al., im Druck). Daher mangelt es an Large-Scale Daten zur Schreibkompetenz in mehreren Sprachen und den zugehörigen empirischen Ansätzen zur Modellierung (Schissel et al., 2019).

Auf Theorieebene gibt es hingegen elaborierte Modelle mehrsprachigen Schreibens. Beispielsweise hat Canagarajah (2015), im Rahmen des translingual approach to writing competence, mehrsprachige Schreibkompetenz als sprachübergreifende Kompetenz konzipiert, die über Fähigkeiten in einzelnen Sprachen hinausgeht. Dieser Ansatz betrachtet Schreibfähigkeiten in unterschiedlichen Sprachen nicht als getrennt, sondern geht von komplexen und dynamischen Beziehungen zwischen den Sprachen im mehrsprachigen Repertoire aus. Auf der Grundlage dieser Theorie kann das mehrsprachige Schreiben als ein mehrdimensionales Konstrukt operationalisiert werden, das sowohl die Beherrschung die einzelsprachlichen Fähigkeiten als auch die sprachübergreifende Schreibkompetenz umfasst (Francis 2012; Cummins, 1979). Bislang wurde diese theoretische Annahme eines integrierten Konstrukts mehrsprachigen Schreibens empirisch nicht modelliert.

Fragestellung

Lassen sich die theoretischen Annahmen zu einem integrierten Konstrukt mehrsprachigen Schreibens in ein empirisches Messmodell übersetzten, das die Konstruktebenen „sprachspezifische Fähigkeiten (proficiency)“ und „sprachübergreifende Kompetenz (competence)“ umfasst? Und bewährt sich dieses Modell an den Daten?

Methode

Für die empirische Modellierung des integrierten Konstrukts mehrsprachigen Schreibens nutzen wir konfirmatorische Faktorenanalysen zweiter Ordnung. Als Daten dienen querschnittlichen Testdaten zum mehrsprachigen Schreiben aus der Studie „Mehrsprachigkeitsentwicklung im Zeitverlauf (MEZ)“ (Gogolin et al. 2017). Unsere Analyse basiert auf einer Stichprobe von n = 968 lebensweltlich mehrsprachigen Sekundarschüler*innen mit Russisch-Deutschen oder Türkisch-Deutschen Sprachhintergrund.

Ergebnisse und ihre Bedeutung

Die Ergebnisse unserer Studie zeigen, dass das von uns entwickelte Messmodell die Zusammenhänge in den empirischen Sprachdaten sehr gut reproduziert. Die vorgebrachten theoretischen Annahmen hinsichtlich eines integrierten Modells mehrsprachigen Schreibens konnten sich somit empirisch bewähren. Die sprachspezifischen Fähigkeitsfaktoren erster Ordnung zeigen substanzielle Zusammenhänge mit den beobachteten Daten (Indikatoren) sowie dem sprachübergreifenden Kompetenzfaktor zweiter Ordnung.

Die Rolle von Lesestrategien in integrierten Schreibaufgaben für Englisch als Fremdsprache

Ximena Delgado-Osorio[1], Johannes Hartig[1], Claudia Harsch[2], Valeriia Koval[2]
[1]DIPF Frankfurt, [2]Universität Bremen

Theoretischer Hintergrund

Akademisches Schreiben erfordert ein breites Spektrum an Kompetenzen, da es eine Kombination aus rezeptiven (Lesen) und produktiven (Schreiben) Fähigkeiten voraussetzt (Chan, 2013). Um die Kombination dieser Fähigkeiten, insbesondere im Kontext von English for academic purposes (EAP) (Hirvela, 2016) zu erfassen, haben integrierte Schreibaufgaben an Popularität gewonnen. Aufgrund des dynamischen Charakters dieser Aufgaben wenden Studierende bei deren Bearbeitung verschiedene zielgerichtete Strategien an, um ihre Fähigkeiten und Handlungen zu koordinieren (Cho

& Lee, 2016; Payant et al., 2019). Plakans (2009) fand, dass Studierende bei integrierten Schreibaufgaben häufig während der gesamten Aufgabenbearbeitung leserelevante Strategien einsetzen und die meisten Leseverstehensstrategien in der Phase vor dem Schreiben verwendet werden.

Da das Leseverstehen die Grundlage von integrierten Aufgaben bildet, stellt sich die Frage, wie die Leistung von L2-Studierenden durch ihre Fähigkeiten und Strategien zum Verstehen des Ausgangstextes beeinflusst wird. In diesem Zusammenhang fanden Plakans & Gebril (2013), dass Teilnehmer*innen mit höherer Leistung besser in der Lage sind, zwischen wichtigeren und weniger wichtigen Ideen in Ausgangstexten zu unterscheiden, um diese in ihren Text einbauen. Daher ist es wichtig, die Fähigkeit von Leser*innen, ihre Aufmerksamkeit während des Lesens zum richtigen Zeitpunkt auf die relevanten Teile des Ausgangstextes zu richten, zu berücksichtigen (Gaddy et al., 2001).

Um das Ziel der Identifkation und Auswahl relevanter Informationen aus dem Ausgangstext zu erreichen, werden häufig Annotationsstrategien wie Hervorhebungen und Notizen eingesetzt. Das Hervorheben gilt hierbei als die beliebteste und als nützlich anerkannte Anmerkungsfunktion (Fariza et al., 2013; Kobayashi, 2009). Wenn Studierende Textinformationen hervorheben und sich hierzu Notizen machen, kann dies das Verständnis verbessern (Gier et al., 2009). Allerdings stellen Hervorhebungen und Notizen wahrscheinlich nicht für alle Fälle oder Situationen die beste Strategie dar (Gier et al., 2010; Bui & Myerson, 2014). Hinsichtlich der genauen Rolle von Annotationstrategien in integrierten Schreibaufgaben, vor allem im Kontext von Englisch als Fremdsprache, besteht daher noch Forschungsbedarf.

Fragestellung

Das Hauptziel dieses Beitrags ist die Untersuchung der Rolle von Annotationsstrategien in integrierten Schreibaufgaben für Englisch als Fremdsprache. Insbesondere interessieren wir uns dafür, wie Annotationsstrategien (Hervorhebung und Notizen) bei der Bearbeitung von integrierten Schreibaufgaben eingesetzt werden und wie sie mit der Schreibleistung der Teilnehmer*innen zusammenhängen.

Methode

Insgesamt wurden von uns vier integrierte Schreibaufgaben entwickelt, welche auf akademischen Einführungstexten aus den Sozial- und Naturwissenschaften beruhen. In zwei dieser Aufgaben wurden die Teilnehmenden darum gebeten, eine Zusammenfassung zu schreiben, während es sich in zwei weiteren Aufgaben um Stellungnahmen handelte. Die vier Stimuli hatten einen ungefähren Umfang von 1000 Wörtern und das erwartete Schreibprodukt sollte zwischen 300-350 Wörtern lang sein. Die Analysen basieren auf einer Stichprobe von Abiturient*innen (n = 116) und Studierenden in englischsprachigen Studiengängen (n = 298). Jede*r Teilnehmer*in bearbeitete in einer Online-Studie eine oder zwei der integrierten Schreibaufgaben, welche per Zufall zugewiesen wurden. Die 674 Schreibprodukte wurden anschließend von sechs geschulten Ratern anhand einer Skala mit neun Analysekriterien, die verschiedenen Textqualitätsdimensionen (Quelltextverwendung, Diskurssynthese und sprachliche Qualität) zugeordnet werden können, bewertet.

Um die Verwendung von Hervorhebungen und Notizen zu analysieren, extrahierten wir die Log-Daten der markierten Textteile und die gemachten Notizen, und kodierten die Verwendung von Hervorhebungen und Notizen entsprechend ihrer Quantität und Qualität. Anschließend führten wir deskriptive Korrelations- und Regressionsanalysen durch.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass bei den meisten Aufgabenbearbeitungen Hervorhebungen vorgenommen wurden (508 Bearbeitungen), während die Notizfunktion nur in der Hälfte der Fälle (208 Bearbeitungen) genutzt wurde. Darüber hinaus zeigten einfache bivariate Korrelationsanalysen, dass der Zusammenhang zwischen Annotationsstrategien und Schreibleistung je nach Aufgabentyp (Zusammenfassungs- und Stellungnahmen-aufgaben) variiert. Weitere Ergebnisse zu Zusammenhängen und Vorhersage der Schreibleistung werden berichtet.

Lesekompetenzmessung in der iKM PLUS

Michael Krelle[1], Jörg Jost[2], Antonia Bachinger[3], Marcel Illetschko[3], Ilka Fladung[2], Sarah Römer[2]
[1]TU Chemnitz, [2]Universität zu Köln, [3]Institut des Bundes für Qualitätssicherung im österreichischen Schulwesen

Theoretischer Hintergrund

Die individuelle Kompetenzmessung PLUS (iKMPLUS) soll als Teil des Pädagogikpakets des BMBWF (2020) zur Unterrichts- und Schulentwicklung und zum Bildungsmonitoring in Österreich beitragen. Dazu werden mit Blick auf die österreichischen Bildungsstandards Kompetenzerhebungen in der 3., 4., 7. Und 8. Schulstufe in verschiedenen Schulformen eingesetzt.

Die iKMPLUS bietet neben dem verpflichtenden Basismodulen eine Reihe an freiwilligen Angeboten zur vertiefenden Lernstandserhebung. So wird eine umfassende Datengrundlage für die pädagogische Arbeit, die Schulentwicklung, das Qualitätsmanagement und das Bildungsmonitoring geschaffen. Eine Reflexion und nachhaltige Verwertung der gewonnenen Informationen sollen so auf unterschiedlichen Ebenen das Schulsystems gezielt angeregt und unterstützt werden.

Im Mittelpunkt des Vortrags stehen die Modellierung und Operationalisierung von zwei wesentlichen Testteilen der iKMPLUS zur Lesekompetenz: Lesefertigkeiten und Leseverstehen. Stellte das Leseverstehen bislang das wissenschaftliche Konstrukt hinter den Kompetenzmessung der iKMPLUS dar, wird mit der Modellierung und Operationalisierung von Lesefertigkeiten im Rahmen der iKMPLUS ein weiterer Kompetenzaspekt einbezogen, der bislang in den iKMPLUS (wie auch in den VERA-Testungen in Deutschland) nicht berücksichtigt wurde.

Theoretisch basieren die Tests auf anderen etablierten Instrumenten (vgl. Auer, Gruber, Mayringer & Wimmer, 2005, Lenhard, Lenhard & Schneider, 2017, Bremerich-Vos & Böhme, 2009, Krelle, 2014), die hier aber mit Blick auf die besonderen Anforderungen der iKMPLUS weiterentwickelt werden. Dazu zählt im Bereich des Leseverstehens die Unterscheidung von hierarchieniedrigen und -hohen Subkompetenzen sowie die Unterscheidung von literarischem Lesen von Epik und Lyrik sowie expositorischem Lesen von kontinuierlichen und diskontinuierlichen Sachtexten. Im Bereich der Lesefertigkeiten werden die Lesegeschwindigkeit und die Lesegenauigkeit als zwei wichtige Aspekte der Leseflüssigkeit getestet.

Fragestellung

Lassen sich aus den neu entwickelten Items zur Messung der Lesefertigkeit und des Leseverstehens reliable Kompetenzskalen bilden und können die jeweiligen Subkompetenzen als eigene Kompetenzfacetten dargestellt werden?

Methode

Grundlage der Untersuchungen ist eine Studie aus 2020 in der Primarstufe (n = 2272 für die 3. Schulstufe und n = 1962 für die 4. Schulstufe für Leseverstehen, n = 476 für Lesefertigkeiten 2022). Die Erfassung der Kompetenzen der Schülerinnen und Schüler erfolgte mit Testaufgaben, die zentrale Bildungsstandards operationalisieren. Mittels Item Response Theorie (IRT, Embretson & Reise 2000 wurden die Leistungen der Schülerinnen und Schüler auf Kompetenzskalen verankert und vergleichbar gemacht.

Ergebnisse und ihre Bedeutung

Es zeigt sich, dass sich die entwickelten Aufgaben gut auf Kompetenzskalen verankern lassen und dass die mit Blick auf das Leseverstehen entwickelte Subkompetenzen eigene Facetten der Lesekompetenz darstellen. Im Bereich der Lesefertigkeiten wird erstmalig im deutschsprachigen Raum eine auf Bildungsstandards bezogene Skala vorgelegt. Die Festlegung von Cut-Scores im Standard-Setting stellt in diesem Bereich eine besondere (auch theoretische) Herausforderung dar, die mit den Teilnehmerinnen und Teilnehmern kritisch diskutiert werden soll. Dabei spielen neben empirischen und fachdidaktischen Erwägungen auch bildungspolitische Setzungen eine Rolle.

B-14: Professionalisierung datengestützter Entscheidungen für Schule und Unterricht

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* T03 R02 D81

Professionalisierung datengestützter Entscheidungen für Schule und Unterricht

Chair(s): **Karina Karst** (Universität Mannheim, Deutschland), **Sarah Bez** (Universität Tübingen/Pädagogische Hochschule Karlsruhe)
*Diskutant*in(nen):* **Hans Anand Pant** (Humboldt-Universität zu Berlin)

Von Lehrpersonen wird zunehmend erwartet, unterrichtliche Entscheidungen auf der Grundlage von Daten zu treffen (Data-based decision-making). Vor dem Hintergrund der Digitalisierung und Datafizierung ist relevant, dass es dafür nicht ausreichend scheint, Daten nur zur Verfügung zu stellen (Hoogland et al., 2016). Als bedeutsame Komponenten auf individueller Ebene gelten hier motivationale Überzeugungen und Kompetenzen bzgl. der Datennutzung (Data Literacy) der entsprechenden Akteure (Filderman et al., 2021).

Datengestützte schulische Entscheidungen werden als zyklisch-sequenzielle Prozesse konzeptualisiert, in denen Daten zielbezogen rezipiert und interpretiert werden müssen, um daraufhin Handlungsentscheidungen treffen, diese umsetzen und bezüglich ihrer Wirkungen evaluieren zu können (Mandinach & Gummer, 2016). Forschungsergebnisse zeigen relativ konsistent, dass Lehrpersonen datengestützte Entscheidungen als herausfordernd wahrnehmen (Hoogland et al., 2016) und eine hinreichend ausgeprägte Data Literacy in den unterschiedlichen Teilkomponenten nicht unbedingt vorausgesetzt werden kann (z.B. Kippers et al., 2018).

Daher wird argumentiert, dass Professionalisierung bezüglich des gesamten Datennutzungszyklus notwendig erscheint, damit dieser gelingen kann und dysfunktionale Wirkungen minimiert werden können (Mandinach & Schildkamp, 2021). Konzeptuelle Arbeiten betonen außerdem, alle Phasen der Lehrer*innenbildung zu adressieren und verschiedene Teilaspekte kontinuierlich anzubahnen (Beck & Nunnaley, 2021). Daher fokussiert das Symposium auf unterschiedliche Professionalisierungsaspekte im Datennutzungszyklus über alle beruflichen Phasen von Lehrpersonen hinweg und adressiert gleichzeitig unterschiedliche individuelle Merkmale von Lehrpersonen, die datengestützte Entscheidungen beeinflussen können.

Beitrag 1 untersucht, wie sich verschiedene Teilkomponenten der Data Literacy über das Studium hinweg entwickeln und wo Studierende besonderen Entwicklungsbedarf aufweisen. Dabei zeigte sich, dass Data Literacy insgesamt erwartungsgemäß zunimmt. Besonders herausfordernd scheint die konkrete Datennutzung sowie die Evaluation der Passung zwischen Anliegen und Datengrundlage.

Beitrag 2 fokussiert, inwiefern die Data Literacy von Lehramtsstudierenden durch eine Intervention zunimmt. Dabei zeigten sich moderate positive Effekte auf Data Literacy, wobei jedoch motivationale Überzeugungen bezüglich Datennutzung weitgehend stabil blieben.

Beitrag 3 widmet sich in einem längsschnittlichen Design der Frage, welche Merkmale datengestützte Unterrichtsentwicklung von Referendar*innen beeinflussen. Hier zeigte sich die Wichtigkeit motivationaler Überzeugungen, während diagnostisches Wissen sowie kollegiale Unterstützung keinen Einfluss zeigte.

Beitrag 4 untersucht inhaltsanalytisch auf der Basis von Reflexionsbögen zu Schüler:innenfeedback, wie konkret die auf diese Datenquelle bezogenen konstruierten Anschlusshandlungen von

Lehrpersonen sind und worauf sie sich beziehen. Dabei leiteten gut 25% der Lehrpersonen konkrete Maßnahmen ab. Dies verweist darauf, dass Informationen aus Daten nur selten in sinnvolle und konkrete Maßnahmen transferiert werden konnten.

Die Beiträge werden abschließend neben sich anschließender Forschungsfragen insbesondere aus einer transferorientierten Perspektive diskutiert, um die Erkenntnisse mit dem Fokus auf praktische Implikationen für die Lehrkräfteaus- und weiterbildung zu synthetisieren.

Beiträge des Symposiums

Was gelingt Studierenden im Bereich Data Literacy leicht, was fällt schwerer? Analyse eines Tests zur Data Literacy von Lehramtsstudierenden

Maurice Wendel, **Karina Karst**
Universität Mannheim

Um das Potential von Daten für die Unterrichtsentwicklung nutzen zu können, benötigen die Lehrkräfte die entsprechenden Kompetenzen im Umgang mit Daten, welche international auch als Data Literacy (DL) gefasst wird (Groß-Ophoff, Koch, Hosenfeld & Helmke, 2006). Eine Studie von Koch (2013) zeigt, dass berufstätige Lehrer*innen in diesem Bereich niedrige Kompetenzen aufweisen. Studierende sollten daher im Rahmen des Lehramtsstudiums entsprechend gefördert werden, um auf die neuen Anforderungen des Lehrerberufs vorbereitet zu sein.

Trotz der zunehmenden Relevanz von Data Literacy für Lehrkräfte existieren in Deutschland noch keine verpflichtenden Angebote zur Förderung ebendieser Kompetenz im Lehramtsstudium (Grillenberger & Romeike, 2019). Zudem wurde noch nicht untersucht, ob sich diese Kompetenz durch das bestehende Lehrangebot im Laufe des Lehramtsstudiums entwickelt. Forschungsfrage 1 untersucht daher, ob DL von Lehramtsstudent*innen im Laufe des Studiums zunimmt. DL für Lehrkräfte wird in diesem Beitrag als ein zirkulärer Prozess im Umgang mit Daten für die Unterrichtsentwicklung definiert, welcher sich aus sechs Kompetenzfacetten zusammensetzt: Data Location, Data Comprehension, Data Interpretation, Data Use, Data Evaluation und Question Posing. Ausgehend von Mandinach & Gummer (2016) und Means, Chen, DeBarger & Pasilla (2011) werden diese Kompetenzfacetten operationalisiert. Demnach wird Data Interpretation unter anderem durch den Indikator „Untersucht die Leistungsheterogenität der Klasse“ beschrieben. Um entsprechende Lehrangebote bedarfsorientiert gestalten zu können, ist es relevant zu verstehen, welche Kompetenzfacetten für Studierende besonders anspruchsvoll sind. Bislang ist unklar, in welchen Bereichen von DL die Lehramtsstudierenden noch Unterstützung durch das Studium benötigen. Mit Forschungsfrage 2 soll daher untersucht werden, wie sich die Kompetenzfacetten von DL hinsichtlich der Ausprägung der Itemschwierigkeiten unterscheiden.

Die Forschungsfragen wurden im Rahmen einer Querschnittstudie an einer baden-württembergischen Universität untersucht. Die Stichprobe besteht aus 200 Lehramtsstudent*innen, von denen 153 im Bachelor und 47 im Master studieren. Zur Messung der Data Literacy wurde ein Test bestehend aus 33 Testitems entwickelt, welche die sechs Kompetenzfacetten von DL abdecken. Der Test wurde nach dem dichotomen Rasch-Modell ausgewertet. Die Personenparameter θ und Itemparameter β werden mit dem bedingten Maximum-Likelihood-Ansatz geschätzt. Die Outfit Kennwerte (outfitMSQ) der Items liegen im Wertebereich von 0.78 bis 1.26, die Infit Kennwerte (infitMSQ) liegen zwischen

0.88 und 1.07. Der DL Test weist eine Andrich's Reliabilität von R = 0.776 auf. Die Verletzung des Differential Item Functioning (DIF) wurde für die Teilungskriterien Geschlecht, Leistung und Studiengang (Bachelor oder Master) geprüft. Es liegen keine Modellverletzungen des DIF vor.

Mittels t-Test zeigt sich, dass die Data Literacy von Bachelorstudent*innen (M = -0.102, SD = 0.881) im Vergleich zu den Masterstudent*innen (M = 0.271, SD = 0.865) signifikant niedriger ausgeprägt ist, t (198) = - 2.553, p = .006, d = 0.426. Zur Analyse der Itemschwierigkeiten werden die Itemparameter (β) des Rasch-Modells nach den gemessenen Kompetenzfacetten gruppiert und analysiert. Die Ergebnisse zeigen die folgenden mittleren Itemschwierigkeiten der Kompetenzfacetten: Data Location (M = -2,239, SD = 1.727), Question Posing (M = -0.513, SD = 0.958), Data Interpretation (M = -0.246, SD = 0.866), Data Comprehension (M = 0.289, SD = 1.111), Data Use (M= 0.337, SD = 1.449) und Data Evaluation (M = 0.689, SD = 1.437).

Die Ergebnisse der Studie weisen darauf hin, dass die Data Literacy im Laufe des Lehramtsstudiums zunimmt. Diese Annahme sollte an einer Längsschnittstudie geprüft werden. Aus der Analyse der Itemschwierigkeiten geht hervor, dass im Mittel die Kompetenzfacetten Data Evaluation und Data Use für die Studierenden am schwierigsten zu lösen sind. Lehrangebote und Interventionen zur Förderung der DL sollten insbesondere diese Kompetenzfacetten adressieren. Weitere Analysen zu den Itemschwierigkeiten folgen und sollen Aufschluss darüber geben, welche inhaltlichen Aspekte für Lehreramtsstudent*innen im Umgang mit Daten besonders herausfordernd sind.

Fördert eine Intervention zu Data Literacy auch motivationale Überzeugungen zu datengestützten Entscheidungen? Ergebnisse einer Studie zu Data Literacy bei Lehramtsstudierenden

Sarah Bez[1], Sebastian Wurster[2], Samuel Merk[3]
[1]Universität Tübingen/Pädagogische Hochschule Karlsruhe, [2]Ludwigs Maximilian Universität München, [3]Pädagogische Hochschule Karlsruhe

Eine hinreichend ausgeprägte Data Literacy sowie positive motivationale Überzeugungen zur Arbeit mit Daten gelten als zentrale Voraussetzungen für adäquate datengestützte Entscheidungen von Lehrpersonen (Hoogland et al., 2016). Allerdings werden datengestützte Entscheidungen von Lehrpersonen als herausfordernd wahrgenommen (Mandinach & Schildkamp, 2021) und Studien zeigen eine eher gering ausgeprägte Data Literacy bei Lehrpersonen (Kippers et al., 2018). Vor diesem Hintergrund wird gefordert, Data Literacy und motivationale Überzeugungen bezüglich der Datennutzung kontinuierlich über alle Phasen der Lehrer*innenbildung hinweg beginnend mit dem Lehramtsstudium zu adressieren, was bisher jedoch nicht systematisch der Fall ist (Beck & Nunnaley, 2021; Mandinach & Gummer, 2016). Erste Interventionsstudien zeigen positive Effekte auf Data Literacy bei Lehramtsstudierenden (Merk et al., 2020; Reeves & Honig, 2015). Bezüglich motivationaler Überzeugungen gegenüber der Datennutzung (konzeptualisiert auf der Basis von Erwartungs-mal-Wert-Modellen) zeigen sich positive signifikante Ergebnisse bei der Selbstwirksamkeitskomponente (z.B. Reeves & Chiang, 2017, 2019), bei den Wert-Komponenten sind die meisten Ergebnisse inkonklusiv (z.B. Reeves & Chiang, 2019; Thoren et al., 2020). Vor diesem Hintergrund untersucht der vorliegende Beitrag Effekte einer Intervention auf Data Literacy (Forschungsfrage 1) und motivationale Überzeugungen bezüglich Datennutzung von Lehramtsstudierenden (Forschungsfrage 2).

Die Intervention orientierte sich an konzeptuellen Arbeiten zu Data Literacy von Mandinach & Gummer (2016) und fokussierte dabei auf die Subfacette Data Use, weil diese bereits sinnvoll im Studium adressiert werden kann (Beck & Nunnaley, 2021; Merk et al., 2020). Die Studie wurde online als Randomized Control Trial (Wartekontrollgruppendesign) durchgeführt. An ihr nahmen N = 136 Lehramtsstudierende einer großen deutschen Universität in Deutschland teil. Die meisten Studierenden befanden sich am Ende ihres Bachelors (MWsemester = 5.5, SDsemester = 1.4), 63% bezeichneten sich als weiblich und 61 % studierten kein MINT-Fach. Zur Erfassung der motivationalen Überzeugungen wurde ein Instrument von Thoren et al. (2020) eingesetzt, da es Eccles und Wigfields (2002) Erwartungs-mal-Wert-Komponenten operationalisiert und diese gleichzeitig differenziert für verschiedene Facetten der Datennutzung erfasst (z.B. Dateninterpretation, Konstruktion von Anschlusshandlungen). Zur Erfassung der Data Literacy wurde ein Test von Merk et al. (2020) eingesetzt. Beide Instrumente zeigten zu beiden Messzeitpunkten ausreichende bis gute Reliabilitäten (motivationale Überzeugungen $\omega >$. 79; Data Literacy: $\alpha >$ .72). Die Effekte der Intervention wurden mit multi-group latent change score-Modellen (Kievit et al., 2018) modelliert und Bayesianisch für das Testen von Hypothesen bei Strukturgleichungsmodellen (Gu et al., 2019) spezifiziert.

Bezüglich Forschungsfrage 1 (Effekt der Intervention auf Data Literacy Test Score) zeigte sich ein substanzieller positiver Effekt (standardisierter Latent Change Score LCS Δ = .76) sowie Evidenz für die Hypothese, dass der Latent Change Score in der Kontrollgruppe gleich 0 und kleiner als in der Interventionsgruppe ist (Posterior Model Probability PMP = .82). Bezüglich motivationaler Überzeugungen zeigte sich Evidenz für positive Effekte auf die Selbstwirksamkeitserwartung in beiden Gruppen für alle Facetten datengestützter Entscheidungen. Diese waren in der Interventionsgruppe größer (LCS $0.31 \leq \Delta \leq 0.72$, PMP = .99). Zudem stiegen die Einschätzungen der Nützlichkeitskomponente in der Treatmentgruppe für die Facette der Formulierung von Fragen und die konkrete Datennutzung an (LCS Δ = .47, PMP = .70), während die anderen Komponenten der motivationalen Überzeugungen bezogen auf alle Facetten datengestützter Entscheidungen relativ stabil blieben.

Insgesamt zeigte sich ein positiver Effekt der Intervention auf die Data Literacy der Lehramtsstudierenden. Die Lehramtsstudierenden berichteten insgesamt positive motivationale Überzeugungen bezüglich datengestützter Entscheidungen, die jedoch mit Ausnahme der Selbstwirksamkeitskomponente und Aspekten der Nützlichkeitskomponente über die Zeit stabil waren. Allerdings muss beachtet werden, dass sich der Inhalt der Intervention hauptsächlich auf die Förderung von Data Literacy fokussierte und weniger die motivationalen Überzeugungen adressierte. Zudem bleibt aufgrund einer nicht vorhandenen Follow-up-Messung offen, wie nachhaltig die Effekte im Professionalisierungsprozess von (angehenden) Lehrpersonen sind.

Längsschnittliche Zusammenhänge zwischen Motivation, Wissen und datenbasierter Unterrichtsentwicklung im Vorbereitungsdienst: Welche Rolle spielt die Ausbildungslehrkraft?

Patrick Hawlitschek, Sofie Henschel
Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin

Theoretischer Hintergrund und Fragestellungen

Datenbasierte Unterrichtsentwicklung umfasst einen komplexen Prozess der systematischen Nutzung schulischer Daten zur Verbesserung der Unterrichtsqualität (Mandinach & Schildkamp, 2021). Vor dem Hintergrund einer steigenden Evidenzorientierung im Bildungswesen gilt der Einbezug von Daten in die Unterrichtsplanung (z.B. Ergebnisse aus Lernstandserhebungen) als wichtiges Element des professionellen Handelns von Lehrkräften (Groß-Ophoff & Cramer, 2022). Jedoch ist die zielgerichtete Datennutzung selbst für erfahrene Lehrkräfte äußerst voraussetzungsvoll und bedarf geeigneter Unterstützungsangebote (Keuning et al., 2017). Demnach scheint es wichtig, bereits Lehrkräfte beim Berufseinstieg in dieser Aufgabe zu unterstützen, weil sie überwiegend keine praktische Erfahrung mit Fremd- und Selbstevaluation besitzen. Dies könnte zu einer nachhaltigen Datennutzung in der Schulpraxis beitragen (Reeves et al., 2016).

Um Datennutzung zu stärken, konzentriert sich die Forschung in Anlehnung an theoretische Rahmenmodelle auf Lehrkräftemerkmale als auch auf schulische Kontextmerkmale (z.B. Helmke & Hosenfeld, 2005). Bisherige Studien zeigten, dass intrinsisch-motivationale Orientierungen und Facetten des pädagogischen Wissens förderlich sein könnten, um Daten systematisch einzubeziehen (Hoogland et al., 2016). Jedoch untersuchten diese Studien förderliche Merkmale primär bei erfahrenen Lehrkräften und fast ausschließlich im Querschnitt. Inwieweit die Lehrkräftemotivation und das Wissen von Lehramtsanwärter*innen die datenbasierte Unterrichtsentwicklung im Verlauf des Vorbereitungsdienstes vorhersagen ist bislang unklar und wird deshalb in diesem Beitrag untersucht (Forschungsfrage 1). Dabei betrachten wir das Wissen in den für die Datennutzung relevanten Facetten „Diagnostik/Evaluation“ sowie „Unterrichtsgestaltung“ (Mandinach & Gummer, 2016). Als Indikator für die intrinsische Lehrkräftemotivation beziehen wir den Unterrichtsenthusiasmus (d.h., die erlebte Freude und Begeisterung für das Unterrichten) ein (Kunter et al., 2011). Für beide Merkmale wurde gezeigt, dass sie positiv mit qualitätsvollem Unterrichten zusammenhängen (z.B. Lazarides et al. (2019), Lohse-Bossenz et al. (2015)). Zudem gibt es Hinweise, dass sie mit der Datennutzung positiv assoziiert sind (Hoogland et al., 2016).

Neben diesen Individualmerkmalen gilt die kollegiale Unterstützung an Schulen als förderliche Ressource für erfolgreiche Datennutzung. Deshalb versuchen Fortbildungsangebote, kollegiale Strukturen unter Lehrkräften zu stärken (z.B. Schildkamp et al., 2016). Für angehende Lehrkräfte gibt es institutionalisiert die Ausbildungslehrkraft, die die Lehramtsanwärter*innen mentoriert und die sie in der datenbasierten Unterrichtsentwicklung unterstützen könnte. Inwieweit die Ausbildungslehrkraft diese Rolle übernimmt, ist noch kein Forschungsgegenstand. Daher untersuchen wir, ob die kollegiale Unterstützung durch die Ausbildungslehrkraft als Moderator fungiert, d.h. ob sie die Effekte der Individualmerkmale verstärkt bzw. abschwächt (Forschungsfrage 2).

Methode

Wir analysieren Längsschnittdaten einer für Nordrhein-Westphalen repräsentativen Stichprobe im BilWiss-Projekt (Kunter et al., 2017). Diese umfasst 670 Referendar*innen (27.19 (SD = 4.20) Jahre, 71.50% weiblich) aus 55 Seminargruppen am Anfang (MZP1) und Ende ihres Vorbereitungsdienstes (MZP2). Die datenbasierte Unterrichtsentwicklung wurde als abhängige Variable zu MZP2 erfasst (Abs et al. (2005); 5 Items; ωMZP2 = .73, Beispiel: „Ich nutze Ergebnisse aus Lernstandserhebungen/

Vergleichsarbeiten zur Überprüfung der Qualität meiner Arbeit“). Die Erfassung des Wissens über „Diagnostik/Evaluation“ (EAP/PV-ReliabilitätMZP1/MZP2 = .72/.72) und über „Unterrichtsgestaltung“ (EAP/PV-ReliabilitätMZP1/MZP2 = .73/.73) erfolgte mit dem BilWiss-Test (Kunter et al., 2017). Der Unterrichtsenthusiasmus wurde nach Kunter et al. (2011; 4 Items; ωMZP1/MZP2 = .80/.80) und die Unterstützung der Ausbildungslehrkraft nach Richter et al. (2011) operationalisiert (12 Items, ωMZP2 = .85). Fehlende Werte (rund 20 %) wurden mit dem R-Paket mice multipel imputiert (m = 20) und mit latenten Strukturgleichungsmodelle ausgewertet. Alle Parameterschätzungen wurden mittels Rubins Regel gepoolt und die Standardfehler korrigiert.

Ergebnisse und ihre Bedeutung

Der Unterrichtsenthusiasmus – aber nicht das pädagogische Wissen – erwies sich als signifikanter, moderat ausgeprägter Prädiktor der zwei Jahre später berichteten, datenbasierten Unterrichtsentwicklung. Demnach entwickeln Lehrkräfte, die Freude beim Unterrichten erleben, ihren Unterricht verstärkt mit Daten. Jedoch lieferten Moderatoranalysen keine Hinweise darauf, dass die Unterstützung seitens der Ausbildungslehrkraft diese Zusammenhänge moderiert. Wir diskutieren die Bedeutung der Lehrkräftemotivation und ob die seitens der Ausbildungslehrkraft vermittelte Unterstützung den Einbezug schulischer Daten bislang nicht adressiert oder die vermittelten Strategien wenig wirksam sind.

Reflexion von Schülerfeedback durch Lehrpersonen

Katrin Kunz, Sebastian Röhl
Universität Tübingen

Rückmeldungen von Schüler*innen zum Unterricht gelten als wichtige Informationsquelle für Lehrpersonen zur Unterrichtsentwicklung und Professionalisierung (Ditton & Arnold, 2004; Zierer & Wisniewski, 2019). Hierfür werden häufig Online- oder Papierfragebögen verwendet, die verschiedene Unterrichtsqualitätsdimensionen erheben. Die so erhobene Information zum Unterricht muss von den Lehrpersonen rezipiert, interpretiert, akzeptiert und reflektiert werden, um mögliche Verbesserungsbereiche aber auch gut funktionierende Unterrichtsaspekte zu identifizieren. Folglich können daraus unterrichtsverbessernde Maßnahmen entwickelt werden (London & Smither, 2002; Röhl et al., 2021), für deren Umsetzung insbesondere der Setzung von konkreten Zielen eine wichtige Rolle zuzukommen scheint (Röhl & Rollett, 2021).

Eine Studie zur Nutzung von Schüler*innenfeedback in den Praxisphasen des Lehramtsstudiums (Neuber & Göbel, 2021) zeigte, dass in der Reflexion des Feedbacks insbesondere die Unterrichtsbereiche Klassenführung, Klarheit und Strukturiertheit sowie das Unterrichtsklima aufgrund von positiven wie kritischen Rückmeldungen eine zentrale Rolle spielten. In Stufenmodellen zur Unterrichtsqualität werden diese Bereiche einerseits als grundlegend für einen effektiven Unterricht, andererseits auch als einfacher umzusetzen beschrieben als beispielsweise Individualisierung und die Förderung selbstgesteuerten Lernens (van der Lans et al., 2019). Für fertig ausgebildete Lehrpersonen liegen bisher keine Befunde vor, welche Bereiche in der Reflexion besonders fokussiert werden und in welchem Maße hieraus eine konkrete Zielsetzung erfolgt. Daher verfolgt dieser Beitrag die folgenden Forschungsfragen:

1. In welchen Unterrichtsbereichen identifizieren Lehrpersonen auf Grundlage von Schüler*innenfragebögen zur Unterrichtsqualität besonders häufig rückgemeldete Stärken und Verbesserungsbereiche?

2. In welchem Maße werden hierdurch Verbesserungsmaßnahmen geplant und weisen diese konkrete Handlungsziele auf?

3. Zeigen sich Unterschiede hinsichtlich der geplanten Verbesserungsmaßnahmen hinsichtlich unterschiedlicher Unterrichtsbereiche?

Im Rahmen eines mehrjährigen Schulversuchs an einem Schulverbund mit insgesamt 1500 Schüler*innen wurde die verpflichtende Einholung eines unterrichtsbezogenen Feedbacks von den Schüler*innen Mithilfe von Unterrichtsqualitätsfragebögen eingeführt. Die Lehrpersonen füllten nach Erhalt des Feedbacks einen vorgegebenen Reflexionsbogen aus. In diesen Bogen trugen die Lehrpersonen die rückgemeldeten Stärken, identifizierten Verbesserungsbereiche und daraus abgeleitete Verbesserungsmaßnahmen ein. Die Auswertung dieser Eintragungen erfolgte mittels der qualitativen Inhaltsanalyse (Kuckartz, 2016). Zur Kodierung der wahrgenommenen Stärken und Verbesserungsbereiche wurde auf Grundlage der in den Feedback-Erhebungsinstrumenten erfragten Unterrichtsqualitätsbereichen ein Kategoriensystem entwickelt. Weiterhin wurden die abgeleiteten Verbesserungsmaßnahmen hinsichtlich deren konkreten und zielorientierten Umsetzbarkeit kategorisiert (Keine; unkonkrete Verbesserungsmaßnahmen, z.B. „Mehr darauf achten“; konkrete Verbesserungsmaßnahmen, z.B. „Erwartungshorizont vor der nächsten Klassenarbeit schriftlich austeilen und besprechen“). Die vorhandenen Eintragungen wurden durch zwei unabhängige Rater kodiert ($.73<\kappa<.90$).

Insgesamt konnten 105 Reflexionsbögen von Schülerrückmeldungen analysiert werden, die von Lehrpersonen am Gymnasium (41.9%), der Real- (22.9%) und Werkrealschule (21.0%) sowie aus der Grundschule (5.7%, fehlend: 8.6%) in den Klassenstufen 3-12 ($M=7.71$, $SD=2.51$) eingeholt wurden.

Besonders häufig gaben die Lehrpersonen an, dass sie Stärken in den Bereichen Unterrichtsklima (77,1%), Klassenmanagement (36,2%) und klare Vermittlung von Unterrichtsinhalten (33,8%) rückgemeldet bekommen haben. Als mögliche Verbesserungsbereiche wurden besonders häufig das Klassenmanagement (38,1%), ein stärkeres Eingehen auf die individuellen Bedürfnisse einzelner Schüler:innen (30,5%) und einer angemesseneren und abwechslungsreicheren Methodik (23,8%) genannt. 33.3% der Reflexionsbögen enthielten keine geplanten Verbesserungsmaßnahmen, weitere 38.1% nur unkonkrete und 28.1% konkrete Verbesserungsmaßnahmen.

Zur Untersuchung der dritten Forschungsfrage wurden X2-Tests hinsichtlich der Konkretheit der Zielgerichtetheit der abgeleiteten Maßnahmen für die drei am häufigsten genannten Verbesserungsbereiche durchgeführt. Es ergibt sich nur für das Eingehen auf individuelle Bedürfnisse eine tendenziell geringere Nennung konkreter Maßnahmen ($X2(2)=5.76$, $p=.056$).

Voll ausgebildete Lehrpersonen scheinen, ähnlich wie Lehramtsstudierende im Praxissemester (vgl. Neuber & Göbel, 2021), besonders häufig Rückmeldungen zu Stärken aus solchen Unterrichtsdimensionen zu erhalten, die in Stufenmodellen zur Unterrichtsqualität als grundlegend und einfacher zu realisieren beschrieben werden. Als Verbesserungsbereiche werden neben dem

Klassenmanagement auch Aspekte der Methodik und individuellen Anpassung identifiziert, die als anspruchsvoller angesehen werden. Bedenklich erscheint der Befund, dass nur etwas mehr als ein Viertel der Lehrpersonen konkrete Verbesserungsmaßnahmen entwickelten. Mögliche Implikationen für die Unterstützung von Lehrpersonen bei der Nutzung von Schülerfeedback werden im Vortrag diskutiert.

B-15: Forschungsdaten erheben, sichern, weitergeben und nachnutzen – Praxisbeispiele aus dem Verbund Forschungsdaten Bildung

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* T03 R03 D75

Forschungsdaten erheben, sichern, weitergeben und nachnutzen – Praxisbeispiele aus dem Verbund Forschungsdaten Bildung

Chair(s): **Sonja Bayer** (DIPF | Leibniz Institut für Bildungsforschung und Bildungsinformation, Germany), **Malte Jansen** (Institut zur Qualitätsentwicklung im Bildungswesen)

Forschung zur Nutzung verfügbarer Daten zeigt, dass die Wirksamkeit von Forschung und Reproduzierbarkeit wissenschaftlicher Ergebnisse eng mit der gemeinsamen Nutzung und Wiederverwendung von Forschungsdaten verbunden sind (Gregory, Cousiin & Groth, 2020). Zudem fördert der Zugang zu FAIRen Forschungsdaten die Transparenz und eine stärkere Ausschöpfung des in den Daten vorhandenen Forschungspotenzials (Nosek et al., 2015, van der Zee & Reich 2018; Pontika et al. 2015; Raffaghelli & Manca, 2019; Bainter & Curran, 2015). Es gehört daher zur guten wissenschaftlichen Praxis Daten zu sichern und nach Möglichkeit zu teilen und nachzunutzen (DFG, 2019; DGfE, GEBF & GFD, 2020). Für Forschende ist das Teilen von Daten jedoch mit Unsicherheiten verbunden, bspw. bezüglich rechtlich-ethischer Aspekte, der Datenaufbereitung und bezüglich technischer Lösungen. Forschende, die Daten nachnutzen wollen, hingegen benötigen Informationen, wie sie geeignete Datensätze finden und Zugriff erhalten. Ziel des Symposiums ist es daher, häufige Fragen Forschender bei der Datenerhebung, -sicherung, -weitergabe und der Nachnutzung aufzugreifen, Informationen zu liefern und gemeinsam zu diskutieren. Der Fokus liegt dabei auf den Angeboten des „Verbunds Forschungsdaten Bildung“ (VerbundFDB).

Der erste Vortrag (Eisentraut) diskutiert rechtliche Aspekte bei der Erhebung, Speicherung, Weitergabe und (Nach-)Nutzung von Forschungsdaten. Hierbei wird bspw. darauf eingegangen, welche datenschutzrechtlichen Aspekte bei der Planung einer Studie und der Erhebung von Forschungsdaten zu beachten sind. Dazu gehören unter anderem die Erstellung einer geeigneten informierten Einwilligung oder auch angemessene Anonymisierungsstrategien. Im zweiten Vortrag (Neuendorf) steht die Sicherung von quantitativen Forschungsdaten und damit zusammenhängend die Möglichkeit ihrer Weitergabe an ein Forschungsdatenzentrum im Mittelpunkt. Es wird ein Überblick darüber gegeben, welche Möglichkeiten zur Archivierung von Forschungsdaten zur Verfügung stehen (z. B. Repositorien vs. Forschungsdatenzentren), Hinweise für die Aufbereitung der Daten gegeben und exemplarisch die Praxis der Datenübergabe an den VerbundFDB dargestellt. Im dritten Vortrag (Jansen) wird die Nachnutzung von Forschungsdaten für Sekundäranalysen thematisiert. Dazu zählen typische Zugangswege sowie Bedingungen der Nachnutzung, die einen Interessensausgleich zwischen Primär- und Sekundärforschenden herstellen sollen. Abschließend wird im vierten Vortrag (Lösch et al.) auf Spezifika der Archivierung und Nachnutzung von qualitativen Datenbeständen wie Audio- und Videodateien oder Interviewtranskripten eingegangen, die einerseits hohes Nachnutzungspotenzial aufweisen, andererseits aber auch meist schwieriger anonymisierbar sind als quantitative Datenbestände.

Beiträge des Symposiums

Forschungsdaten erheben – Rechtliche Voraussetzungen zur (Nach-)Nutzung von Forschungsdaten

Marcus Eisentraut
GESIS – Leibniz-Institut für Sozialwissenschaften

Forschungsdaten sind ein zentraler Bestandteil der empirischen Bildungsforschung und ihre Erhebung eine grundlegende Voraussetzung für viele Forschungsprojekte. Dabei kommt es häufig zu Mehrfacherhebungen von vergleichbaren Daten über dieselben Untersuchungsobjekte. Diese Praxis erhöht nicht nur den zeitlichen und monetären Aufwand der Forschenden. Insbesondere dort, wo sogenannte personenbezogene Daten generiert werden, steht die Mehrfacherhebung im Widerspruch zum Grundsatz der Datensparsamkeit und der Vermeidung redundanter Informationssammlungen. Vor diesem Hintergrund fordern heute immer mehr Drittmittelgeber, wie etwa die EU im Rahmen von Horizon 2020 oder das BMBF im Bereich der empirischen Bildungsforschung, dass Forschungsdaten aus geförderten Projekten nachnutzbar gemacht werden sollen.

Neben der Nachvollziehbar- und Verständlichkeit spielt für die Nachnutzung die Gewährleistung unterschiedlicher rechtlicher Voraussetzungen eine wichtige Rolle. Hierzu zählen beispielsweise die Regelungen des Datenschutzes im Umgang mit personenbezogenen Daten, wie Auflagen zum Einholen der informierten Einwilligung, zur Anonymisierung der Daten oder zum geschützten Speichern personenbezogener Informationen.

Der Vortrag adressiert die oben beschriebene Thematik, indem verschieden die rechtlichen Voraussetzungen im Umgang mit Forschungsdaten kurz skizziert werden, um ein grundlegendes Verständnis für die diese Aspekte des Forschungsdatenmanagements zu entwickeln und damit Forschungsdaten rechtskonform auf die Nachnutzung vorzubereiten.

Forschungsdaten sichern und weitergeben

Claudia Neuendorf
Universität Tübingen

Die Sicherung und Weitergabe von Forschungsdaten an ein Forschungsdatenzentrum wird immer häufiger von Mittelgebern gefordert und auch innerhalb der wissenschaftlichen Community gibt es in letzter Zeit zunehmend Leitlinien, die die Weitergabe von Forschungsdaten zum Zwecke der Replikation von Forschungsergebnissen und zur Sekundärnutzung fordern (Allianz der deutschen Wissenschaftsorganisationen, 2010, DFG Fachkollegium Erziehungswissenschaft: Stanat, 2014, Deutsche Gesellschaft für Erziehungswissenschaft, 2017, DGPs: DGPs-Kommission „Open Science“, 2020; Schönbrodt, Gollwitzer & Abele-Brehm, 2016). Für Forschende stellen sich somit die Fragen 1) ob die Daten überhaupt zur Archivierung weitergegeben werden dürfen, 2) ob sie auch zur Nachnutzung bereitgestellt werden sollen und 3) wo und unter welchen Bedingungen sie konkret bereitgestellt werden können.

Im Beitrag werden diese Fragen mit einem Fokus auf die Bildungsforschung behandelt. Zunächst wird darauf eingegangen, welche rechtlichen Voraussetzungen erfüllt sein müssen. Dies betrifft den Schutz personenbezogener Daten, ggf. die Beachtung von Urheberrechten bei der Datendokumentation sowie besonders im Bereich der Erhebung an Schulen mögliche Auflagen von genehmigenden Behörden. Der Verbund Forschungsdaten Bildung (VerbundFDB) hat zu diesen Fragen Strategien und Handreichungen entwickelt, die hier vorgestellt werden sollen (z.B. Meyermann & Porzelt, 2017).

Zur Frage, ob im Forschungsprojekt erhobene Daten bereitgestellt werden sollten, gibt es trotz der zunehmenden Forderung der Verfügbarmachung von Forschungsdaten gewisse Einschränkungen. Dies betrifft in erster Linie die Abwägung von Kosten und potenziellem Nutzen der Bereitstellung der Daten. Zur Bewertung des Nachnutzungspotentials werden erste Ansätze aus dem VerbundFDB vorgestellt.

Weiterhin wird die Suche nach einem passenden Publikationsagenten mithilfe des Verzeichnisses re3data.org vorgestellt und es werden mögliche Qualitätsanforderungen bei der Auswahl eines Repositoriums formuliert. Der VerbundFDB versteht sich als zentrale Anlaufstelle, die Forschungsdaten aus der Bildungsforschung aufnimmt und an diejenigen Datenzentren verteilt, welche die größte Expertise für den entsprechenden Datentyp haben. Das Prinzip des zentralen Zugangs und der verteilten Archivierung wird vorgestellt.

Schließlich wird beispielhaft der Prozess der Abgabe und Bereitstellung von Forschungsdaten durch das Forschungsdatenzentrum am Institut zur Qualitätsentwicklung im Bildungswesen vorgestellt.

Forschungsdaten nachnutzen: Bedingungen, Chancen und Herausforderungen

Malte Jansen
Institut zur Qualitätsentwicklung im Bildungswesen

In der deutschen Bildungsforschungslandschaft werden langsam die Vorteile der Sekundärnutzung von Datensätzen erkannt (Jansen et al., 2021). So wird die Rekrutierung von Bildungsinstitutionen zur Teilnahme an empirischen Studien immer schwieriger, weil diese sich durch die vielen Studien zunehmend belastet fühlen und Datenerhebungen für die Institution und die Teilnehmenden mit teilweise erheblichem Aufwand und häufig auch dem Verlust von Unterrichtszeit verbunden sind. Werden Forschungsdaten der wissenschaftlichen Community für Sekundäranalysen zur Verfügung gestellt, so können vorhandene Datensätze, die mit teils erheblichem personellen und finanziellen Aufwand erhoben wurden, breiter und tiefer ausgewertet werden. Auch kann die Anzahl neuer Erhebungen reduziert werden, wenn bereits qualitativ hochwertige Datensätze zur Beantwortung einer Forschungsfrage vorhanden sind. Schließlich gebieten auch die Grundsätze guter wissenschaftlicher Praxis, dass der wissenschaftlichen Community die Möglichkeit der Replikation und Reanalyse gegeben wird wofür eine Datenweitergabe (zumindest auf Anfrage) unerlässlich ist.

Im geplanten Beitrag sollen zunächst auf allgemeiner Ebene Chancen, aber auch Herausforderungen von sekundäranalytischer Forschung im Bildungsbereich beschrieben werden. Anschließend soll ein Überblick über verschiedene Modelle der Datennachnutzung gegeben werden und deren Vor- und Nachteile gegeben werden (z. B. Datenzugang auf Anfrage über Primärforscherinnen und Primärforscher, offener Datenzugang, Datenzugang auf Antrag). Dabei wird insbesondere auch darauf eingegangen, wie ein Interessensausgleich zwischen Datenproduzierenden und Datennachnutzenden hergestellt werden kann.

Im nächsten Schritt wird die Perspektive der Datennachnutzenden eingenommen. Es wird erläutert, wie relevante Datenbestände aufgefunden werden können und dafür auf mehrere nationale und internationale Suchportale verwiesen. Beispielhaft wird die Suche auf www.Forschungsdaten-Bildung.de genauer beschrieben. Anschließend wird auf typische Zugangsbedingungen eingegangen. Für den Fall eines antragsbezogenen Verfahrens werden typische Bedingungen der

Datennachnutzung beschrieben und exemplarisch das Verfahren der Datenbeantragung am Forschungsdatenzentrum des Instituts zur Qualitätsentwicklung im Bildungswesen (FDZ am IQB) vorgestellt.

Die Spezifika der Erhebung, Sicherung, Weitergabe und Nachnutzung qualitativer Daten

Thomas Lösch, Maike Porzelt, Nadeshda Jung, Doris Bambey, Marius Gerecht
DIPF | Leibniz Institut für Bildungsforschung und Bildungsinformation, Germany

Die Erhebung, Sicherung, Weitergabe und Nachnutzung qualitativer Forschungsdaten in der empirischen Bildungsforschung stellt sowohl an die Datenproduzenten als auch die Forschungsdatenzentren eine Reihe spezifischer Anforderungen. Hohe biografische Anteile am gesammelten Datenmaterial führen zur Kollektion von sensiblen, teils hoch emotionalen Daten, die oft einschneidende Lebensereignisse und Wendungen im Leben der Probanden dokumentieren. Die Preisgabe dieser Daten setzt einen hohen Vertrauensvorschuss seitens der Probanden gegenüber der Forschung und den Forschenden selbst voraus. Darüber hinaus sind auch methodische Grenzen gesetzt, wenn es um die Anonymisierung des Datenmaterials geht, soll doch das analytische Potenzial der Daten nicht durch das Anonymisierungsprozedere gänzlich ausgehöhlt werden. Daher wurden und werden qualitative Daten oft nicht oder nur sehr rudimentär anonymisiert zur Nachnutzung weitergegeben. Die Basis eines vertrauensvollen Umgangs mit Forschungsdaten sind daher transparente, harmonisierte Verfahren sowie eindeutige, vertragliche Vereinbarungen zwischen allen am Datenerhebungs- und auch Datenverwertungsprozess beteiligten Protagonisten. Zudem erfordern viele methodische Zugänge qualitativer Forschung eine tiefgehende Beschreibung der Erhebungssituation und der Daten selbst, um eine adäquate Ausgangsbasis für die Auswertung zu erhalten, was wiederum die Anonymisierung der Daten zusätzlich erschwert. Der Zugriff Dritter auf nicht anonymisierte Daten wiederum stellt spezifische Anforderungen an die vorhandene und sich (weiter)entwickelnde Forschungsdateninfrastruktur. Das FDZ Bildung am DIPF hat Lösungen für die Dokumentation und Weitergabe qualitativer Forschungsdaten entwickelt, die auf die Spezifika der empirischen Bildungsforschung zugeschnitten sind. Dazu zählen transparente und rechtssichere Verfahren der Datenweitergabe und die zielgruppenspezifische Dokumentation der Daten selbst.

In diesem Beitrag wird im Besonderen auf die Informationsdichte und Kontextsensitivität qualitativer Daten eingegangen und was dies für deren Bearbeitung im Forschungsprozess sowie eine mögliche Bereitstellung der Daten zur Nachnutzung bedeutet.

B-16: Designing effective digital learning environments

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* R11 T00 D01

Designing effective digital learning environments

Chair(s): **Andreas Gegenfurtner** (Universität Augsburg, Germany), **Ingo Kollar** (Universität Augsburg)
*Diskutant*in(nen):* **Judith Schweppe** (Universität Passau)

This symposium is dedicated to the design of effective digital learning environments in education. Research on the instructional design of technology-enhanced learning has prospered for many years, producing significant advances related to design principles and theoretical frameworks in numerous digital contexts—for example, with regard to dynamic visualizations, pedagogical agents, gamification, and computer-supported collaborative learning (CSCL). For an individual scholar, it is not always easy to keep track of the myriad of original empirical reports published on different digital learning environments. As a remedy, this symposium seeks to offer a synthesis of research evidence from previous meta-analyses and systematic literature reviews on the predictive validity of particular theoretical arguments on how to design digital environments that support learning, motivation, affect, and collaboration.

Consequently, the aim of this edited volume is to bring together leading experts for sharing their most recent advancements on the design of effective digital learning in a total of four symposium papers. Each paper focuses on another digital technology and the paper authors present a summary of predominantly meta-analytic evidence about effective design principles, their boundary conditions, and implementation recommendations in educational settings. Paper 1 addresses dynamic visualizations in education with reference to the cognitive theory of multimedia learning (Mayer, 2021) and the cognitive-affective-social theory of learning in digital environments (Schneider et al., 2021), highlighting the segmenting principle, the signaling principle, and the inclusion of human instructors in dynamic visualizations as strategies to further improve learning with dynamic visualizations. Paper 2 focuses on design recommendations for using pedagogical agents—fictitious social entities in virtual learning environments—based on cognitive load theory (Sweller et al., 2021), cognitive affective theory of learning with media (Moreno, 2006), the integrated cognitive affective model of learning with multimedia (Plass & Kaplan, 2015), and the cognitive-affective-social theory of learning in digital environments (Schneider et al., 2021). Paper 3 discusses the effectiveness of gamification in education based on meta-analytic and review evidence that draws particularly from Landers' (2015) theory of gamified learning (2015) and Bedwell et al.'s (2012) game element taxonomy. Finally, Paper 4 addresses the effective design of computer-supported collaborative learning grounded in an in-depth review of seven meta-analyses on CSCL that test the assumption that collaborating with others has a strong potential to facilitate learning (Stahl & Hakkarainen, 2021). To conclude this symposium, an internationally renowned expert in the field will discuss the theoretical and educational implications of the symposium contributions.

Beiträge des Symposiums

Dynamic visualizations: How to overcome challenges and seize opportunities

Martin Merkt[1], **Markus Huff**[2]
[1]DIE Bonn, [2]Universität Tübingen

Introduction

This contribution focuses on challenges and opportunities associated with dynamic visualizations as learning materials. We use a broad definition of dynamic visualizations, highlighting the transience of information as their defining feature. Despite the requirement to pay continuous attention to system-paced dynamic visualizations, meta-analyses show the superiority of dynamic visualizations over static visualizations, especially when the learning materials depict changes over time (Höffler & Leutner, 2007; Plötzner et al., 2020). We give an overview of meta-analyses and systematic literature reviews regarding the segmenting principle, the signaling principle, and the inclusion of human instructors in dynamic visualizations as strategies to further improve learning with dynamic visualizations.

Segmenting Principle

Considering learners' limited cognitive resources (Mayer, 2021; Sweller et al., 2019), the transient flow of information in dynamic visualizations increases the risk of cognitive overload (Leahy & Sweller, 2011). Consequently, the Cognitive Theory of Multimedia Learning (CTML; Mayer, 2021) proposes to provide learners with segmented dynamic visualizations that stop at pre-defined breakpoints (segmenting principle). In their meta-analysis, Rey et al. (2019) distinguished between system-paced (i.e., the presentation continues automatically after a break) and learner-paced segmenting (i.e., the presentation continues after learners' input). System-paced segmenting positively affected both retention and transfer, whereas learner-paced segmenting only positively affected transfer. Contrary to theoretical assumptions and previous studies (e.g., Spanjers et al., 2011), Rey et al. (2019) observed larger segmenting effects on retention for participants with high levels of prior knowledge. This finding warrants future research (also see Mayer, 2021).

Signaling Principle

Transient presentation of information constrains learners' time for selecting, organizing, and integrating relevant pieces of information, which are essential prerequisites for learning (CTML, Mayer, 2021). To support these processes in static and dynamic visualizations, Mayer (2021) suggests the signaling principle, meaning that relevant information should be highlighted by verbal (e.g., headings) and/or visual cues (e.g., arrows). Two meta-analyses corroborate the effectiveness of signaling on retention, comprehension, and transfer with small to medium effect sizes (Richter et al., 2016; Schneider et al., 2018). Whereas Schneider et al. (2018) found no evidence for prior knowledge as a moderator, Richter et al. (2016) observed that signaling effects on comprehension and transfer were only evident in learners with low to medium levels of prior knowledge. Interestingly, information transience was not identified as a moderator for signaling effects.

Human Instructors

The Cognitive-Affective-Social Theory of Learning in digital Environments (CASTLE, Schneider et al., 2021) highlights the importance of social cues in the learning process because they are assumed to foster deeper elaborations of the learning materials. Dynamic visualizations allow for the inclusion of strong social cues such as human instructors. Even though a visible human instructor was shown to compete for visual attention with relevant learning materials (Kizilcec et al, 2014; van Wermeskerken et al., 2018), previous research mainly observed neutral or positive effects of human instructors in

dynamic learning materials (Henderson & Schroeder, 2021). However, further research is necessary to address potential boundary conditions to investigate whether potential adverse effects through visual competition are always compensated by positive effects of instructor presence or corresponding learning strategies.

Conclusion

This contribution provides an overview of different design principles for dynamic visualizations. Given that these principles are both based on theory and corroborated in meta-analyses and systematic literature reviews, we are confident that they are helpful in designing dynamic visualizations for learning. At the same time, this contribution also highlights the need for further research to identify potential boundary conditions for the effectiveness of the principles.

Learning with pedagogical agents in digital environments

Maik Beege[1], Sascha Schneider[2], Steve Nebel[3], Günter Daniel Rey[3]
[1]PH Freiburg, [2]Universität Zürich, [3]TU Chemnitz

When instructional interactions are moving from classical classroom environments to digital learning spaces, the social exchange between learner and instructor diminished. This might result in hindered communication and consequently, learning (Burgoon et al., 2000). To counteract this, designers can implement fictitious social entities in learning environments. These virtual characters can be defined as pedagogical agents (PAs). PAs digital characters that serve instructional or conversational purposes in educational settings (Veletsianos & Russel, 2013; Martha & Santoso, 2019).

The effects of PAs on learning processes are discussed regarding general multimedia learning theories. According to the Cognitive Load Theory (CLT; Sweller et al., 1998, 2019), cognitive resources are limited and learning irrelevant load should be minimized. Since PAs tie up cognitive resources, designer should at least compensate this potential negative effect through the use of instructional design principles. For example, we will discuss the supporting effects of providing cognitive guidance or highlighting important information. However, according to the Cognitive Affective Theory of Learning with Media (CATLM; Moreno, 2006) and the Integrated Cognitive Affective Model of Learning with Multimedia (ICALM; Plass & Kaplan, 2015), emotional as well as motivational processing can influence learning in a positive way. Since, PAs are suitable for performing or show authentic emotional cues (Adamo et al., 2021), emotional and motivational impacts of PAs on learning will be further discussed. Finally, the Cognitive-Affective-Social Theory of Learning in digital Environments (CASTLE; Schneider et al., 2021) postulates that activating social schemata in the long-term memory through social cues can foster learning. PAs trigger a social context and consequently, impacts of social cues performed or shown by PAs will be further outlined, considering the effectiveness of PAs on learning processes.

Even if several meta-analyses (Schroeder et al., 2013; Castro-Alonso et al., 2021; Davis et al., 2022) indicate a significant positive effect of the presence of PAs on learning with small effect sizes, we argue that PAs should be included with caution since the concrete design and use of design features might determine if PAs are beneficial for learning or not. Moreno (2005) already pointed out that PAs may be effective for learning due to their pedagogy rather than merely their appearance. Thus, the question should not be whether to use an PA or not, but rather how to design PAs in an instructional

effective manner. This is in line with results of the mentioned meta-analyses that found several moderators for the effect of PAs on learning.

We finally outline that there is a bright scope for future research. There are only a few studies that investigated several instructional roles of PAs. Furthermore, there are still design features regarding cognitive, affective, motivational and social processes that are not thoroughly investigated. In particular design features depending on the target group of the instructional intervention should be considered. The cost-effectiveness of PAs should be further considered by future researchers and designers. Technological advantages allow the pro-duction of various PAs at a relatively low cost. Design effects with low to medium effect sizes do still have high practical importance since production costs lowered and software complexity decreased due to new animation software.

Effectiveness of gamification in education

Ruben Schlag[1], Michael Sailer[2], Daniel Tolks[3], Manuel Ninaus[4], Maximilian Sailer[1]
[1]Universität Passau, [2]LMU München, [3]Universität Bielefeld, [4]Universität Graz

Initially a marketing concept, gamification has been implemented in learning settings involving digital technologies to improve learning processes. Educational practitioners and researchers alike wonder how “the use of game design elements in non-game contexts“ (Deterding et al., 2011, p. 2) can positively influence learning outcomes. Recent meta-analyses showed significant, small to medium positive overall effects of gamification or game elements on learning outcomes (Ritzhaupt et al., 2021). Still it needs more empirical evidence on its effectiveness and limitations in order to respond to critiques of the approach. This presentation at the symposium intends to showcase the use of gamification and its potentials in education by focusing on theoretical frameworks, game elements and empirical evidence. First of all, we want to differentiate between prevalent theoretical concepts used for empirical research on gamification and game-based learning. As Krath and colleagues (2021) identified 118 different theories to be in use, this research illustrates the variety of theories applied in gamification research and the pervasiveness of gamification across various domains. We will highlight Landers’ theory of gamified learning (2015) on this matter. Landers (2015) introduces four components that describe the relationship between gamification and learning: (1) game elements, (2) instructional content, (3) behavior/attitude and (4) learning outcomes by drawing on the game element taxonomy of Bedwell and colleagues (2012). On a next subtopic, we display a number of prominent game elements like “Points” and “Leaderboards” for example that have been commonly used in education and are examined in empirical research. We will shorty outline how these elements might interlink in practice. Furthermore, results of selected meta-analyses are delineated. Lastly, moderating variables are discussed to more closely examine the effects of gamification on learning outcomes. These moderators include factors such as the amount of time gamified interventions last for (Bai et al., 2020), Type of publication (Huang et al. 2020) among others. This presentation wants to illustrate gamification as an effective method of technology-enhanced instruction, while acknowledging its limitations.

What are the effects of computer-supported collaborative learning on learning processes and outcomes? A review of available meta-analytical evidence

Ingo Kollar[1], Martin Greisel[1], Tugce Özbek[1], Laura Spang[1], Freydis Vogel[2]
[1]Universität Augsburg, [2]University of Nottingham

Computer-supported collaborative learning (CSCL) subsumes a broad variety of different learning scenarios and instructional approaches that (a) have students learn in groups (in contrast to learning in individualized settings) and (b) use some kind of digital technology to support learning and collaboration among students. At a theoretical level, collaborating with others has a strong potential to facilitate learning, which is described for example in (socio-)cognitive and sociocultural theories of learning (e.g., Stahl & Hakkarainen, 2021). Digital technologies are designed to enhance this potential for learning, for instance through structuring social learning processes. Yet, researchers discuss, for example, if computer support is really necessary, or if structuring the collaboration has impeding side-effects on students' motivation.

In this overview of meta-analytic evidence for effects of CSCL on learning, we therefore asked four research questions: (1) What are the overall effects of CSCL instructions on learning outcomes when compared to other kinds of instruction? (2) To what extent does the inclusion of computers enhance the effects of collaborative learning? (3) What are the effects of collaboration scripts as a specific form of CSCL instructions? (4) Which learning processes mediate the effects of CSCL on learning?

We performed an in-depth review of meta-analyses on CSCL (N = 7; Chen et al., 2018; Jeong et al., 2019; Radkowitsch et al., 2020; Saqr, et al., 2022; Sung, et al., 2017; Talan, 2021; Vogel et al., 2017). With respect to RQ1, the evidence suggests medium effects of CSCL on learning outcomes when compared to other instruction forms. For example, the meta-analysis by Jeong et al. (2019), based on 316 outcomes from 143 studies with N = 12.380 students which compared CSCL with some kind of control group (ranging from individual learning with or without computers to traditional classrooms using little or no technology) yielded an overall effect size of Hedge's g = 0.51. Concerning RQ2, meta-analytical evidence suggests positive, medium-to-large sized effects of adding computers to collaboration on learning outcomes. Based on results from 71 studies with a total of N = 11.286 students, Chen et al. (2018), for example, found collaborative learning with computers to be significantly more effective than collaborative learning without computers, with effect sizes ranging from Hedge's g = 0.45 (for knowledge acquisition) to g = 0.89 (for group task performance). Concerning RQ3, a meta-analysis by Vogel et al. (2017), based on 22 articles with a total of N = 2.825 participants, found a small positive effect of CSCL scripts compared to unstructured CSCL on the acquisition of domain knowledge (Cohen's d = 0.20), and a large positive effect on the acquisition of cross-domain skills (Cohen's d = 0.95). The meta-analysis by Radkowitsch et al. (2020) adds that CSCL scripts also do not have a negative effect on motivation (Hedge's g = 0.13). Finally, with respect to RQ4, Saqr et al. (2022) aggregated the results of 19 studies with a total of N = 12.293 participants that gathered CSCL process data through Social Network Analysis (SNA) and showed that the centrality indicators under consideration were on average correlated with achievement by r = .46, which indicates that the more learners engage in communication during CSCL, the higher is their achievement.

To conclude, CSCL seems to be an effective way to foster learning, especially when it uses computer technology to provide guidance and when it manages to get learners actively involved in collaboration. Designers and practitioners should thus aim at developing and using CSCL environments that

structure the learning processes in a way that guarantees that all learners take an active role during collaboration.

B-17: Phasen digital-gestützten individualisierten Übens im Fokus: Erkenntnisse aus dem CODIP-Projekt

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* T03 R04 D10

Phasen digital-gestützten individualisierten Übens im Fokus: Erkenntnisse aus dem CODIP-Projekt

Chair(s): **Alina Kristin Hase** (Leuphana Universität Lüneburg, Deutschland), **Poldi Kuhl** (Leuphana Universität Lüneburg, Deutschland)
*Diskutant*in(nen):* **Julia Gerick** (TU Braunschweig, Deutschland)

Ausgehend vom Primat der Pädagogik sollten digitale Medien im Fachunterricht nur dann eingesetzt werden, wenn hiermit ein Mehrwert für das Lernen verbunden ist (Zierer, 2018). In Phasen des Übens werden die Potentiale digitaler Medien besonders deutlich, da sie Lernprozesse individualisiert begleiten und unterstützen können (Gerjets & Scheiter, 2019; Livingstone, 2012; Schaumburg, 2021). Voraussetzung für das digital-gestützte Üben ist die Bereitschaft der Lehrkräfte, entsprechende digitale Medien zu nutzen (Scherer & Teo, 2019). Der Gestaltung digital-gestützten individualisierten Übens widmet sich das CODIP-Projekt, welches im Rahmen der Qualitätsoffensive Lehrerbildung gefördert wird. Das Symposium bietet Einblick in vier CODIP-Teilprojekte, die sich in unterschiedlichen Disziplinen mit digital-gestütztem individualisierten Üben in Unterrichtsforschung und Lehrkräftebildung auseinandersetzen.

Im ersten Beitrag geht es darum, die Faktoren zu identifizieren, die mit der Technologieakzeptanz im Zusammenhang stehen. Dazu wird im Zuge einer systematischen Übersichtsarbeit mit *N*=6 empirischen Studien explizit die Akzeptanz von Lehrkräften gegenüber digitalen Lernplattformen zum Üben betrachtet. Für die Nutzung digitaler Lernplattformen zum Üben scheint vor allem das Bewusstsein über die Nützlichkeit sowie ein hohes Maß bisheriger Erfahrung in der Nutzung von Technologie relevant zu sein.

Der zweite Beitrag widmet sich der digital-gestützten Sportspielvermittlung. Es wurde eine App für die Unterstützung der Reflexionsprozesse der Schüler:innen eingesetzt. Die Studie mit *N*=104 Viertklässler:innen untersucht anhand von leitfadengestützten Interviews, die nach der Grounded Theory Methodologie ausgewertet wurden, die Perspektiven und Sichtweisen der Schüler:innen auf den Einsatz digitaler Medien im Sportspielunterricht. Hier wird deutlich, dass das digitale Medium als Hilfsmittel im Lernprozess gedeutet werden kann und die Schüler:innen den Einsatz von Tablets im Sportunterricht kritisch reflektieren.

Im Rahmen des dritten Beitrags wird aufbauend auf Erkenntnissen der Forschung zu unterrichtlichem Feedback für Lernende die Wahrnehmung und der Einfluss zweier verschiedener Feedbackarten (knowledge of result vs. elaboriert) auf mathematische Lernleistungen in einer Laborstudie betrachtet. Aus der Studie mit *N*=147 Siebtklässler:innen geht hervor, dass die Feedbackarten nicht unterschiedlich wahrgenommen werden. Analysen ergeben die Notwendigkeit der positiven Wahrnehmung elaborierten Feedbacks zur Leistungsentwicklung der Schüler:innen.

Der vierte Beitrag befasst sich mit der Nutzung von Daten, die Lehrkräfte beim Einsatz digitaler Lernplattformen erhalten. Eine Interviewstudie mit *N*=12 Grundschullehrkräften überprüfte auf Grundlage der inhaltlich strukturierenden qualitativen Inhaltsanalyse die Anwendbarkeit des Datennutzungskreislaufes, der im Bereich der datenbasierten Entscheidungsfindung bereits etabliert

ist, aber bislang noch nicht konkret auf Daten aus digitalen Medien angewendet wurde. Hierbei zeigt sich, dass Lehrkräfte digitale Daten bislang größtenteils nur nutzen, um Ergebnisse zu kontrollieren und darüber hinaus nur in geringerem Maße auch zur individualisierten Unterrichtsgestaltung.

Beiträge des Symposiums

Technologieakzeptanz als Voraussetzung zum digital-gestützten Üben – Eine systematische Literaturübersicht

Leonie Kahnbach, Dirk Lehr
Leuphana Universität Lüneburg, Deutschland

Theoretischer Hintergrund

Digitale Lernplattformen habe das Potenzial, das Lernen und Üben zu bereichern, vor allem wenn sie zum Zweck des Individualisierens eingesetzt werden (Schaumburg, 2019; Major, 2021). Sie zeichnen sich dadurch aus, dass Lehrkräfte den Lernfortschritt der SchülerInnen einsehen können und individuelles Feedback und Lernwege zur individuellen Förderung erstellen können. In Deutschland gab es durch die Covid-19 Pandemie zwar einen Digitalisierungsschub (Forsa, 2021) und die Nutzung von Bildungstechnologien gilt besonders zum Zweck der Individualisierung als wünschenswert (Kultusministerkonferenz, 2021), jedoch werden bisher kaum Lernplattformen zu diesem Zweck genutzt. (Forsa, 2020; ICILS, 2018; Mutmaß et al., 2021). Da es in Deutschland jedoch ein rasch wachsendes Angebot von Lernprogrammen zum digital-gestützten, individualisierten Lernen und Üben gibt (Holmes et al., 2018), stellt sich die Frage, welche Faktoren mit der verhältnismäßig geringen Nutzung zusammenhängen. In diesem Kontext präsentiert die Technologieakzeptanzforschung seit einigen Jahren wichtige Erkenntnisse über die Faktoren, die mit der Intention von Lehrkräften, Bildungstechnologien im Unterricht zu nutzen einher gehen. Die Ergebnisse wurden bereits systematisch und meta-analytisch zusammengefasst, jedoch adressierten diese Übersichtsarbeiten bisher vor allem die Aufklärung der Intention zur Nutzung von Technologie im schulischen Kontext zu einem nicht näher spezifizierten Zweck: Scherer und Teo (2019) fassten zuletzt meta-analytisch zusammen, welche Faktoren aus dem Technologie Akzeptanz Modell (TAM; Davis et al., 1989) mit der Intention von Lehrkräften, Bildungstechnologie zu nutzen, im Zusammenhang stehen. Der Großteil der einbezogenen Studien fokussierte sich dabei auf Lehramtsstudierende, sowie auf Technologienutzung im Allgemeinen. Da jedoch die Nutzung von Technologie im schulischen Kontext stets einem lernförderlichen Zweck dienen sollte und auch der Kontext der Technologienutzung damit zusammenhängt, welche Faktoren für die Aufklärung der Intention zur Nutzung relevant werden (Nistor, 2020), erscheint es sinnvoll, die Faktoren zu identifizieren, die die Intention von Lehrkräften, Technologie zu einem lernförderlichen Zweck zu nutzen, bestärken. Eine zusammenfassende Betrachtung solcher Akzeptanzstudien gibt es bislang nicht. Daher hat der vorliegende Beitrag das Ziel, die Erkenntnisse der Studien, welche die Akzeptanz der Nutzung von Technologien zum Zweck des digital-gestützten individualisierten Übens fokussieren, zusammen zu fassen.

Fragestellung

Welche Faktoren stehen mit der Intention von Lehrkräften zur Nutzung von Lernplattformen zum digitalen Üben in Zusammenhang?

Methode

Die systematische Übersicht wurde in Anlehnung an den bevorzugten Report von Items für systematische Übersichten und Meta-Analysen „Preferred Reporting Items for Systematic Reviews and Meta-Analysis“ (PRISMA) in der deutschen Übersetzung von Ziegler, Antes und König (2011) durchgeführt. Die Suche fand im Juli 2022 statt. Nach der Erstellung eines Such-Strings, wurde innerhalb der Datenbank Scopus sowie in sieben fachspezifischen Datenbanken via EbscoHost gesucht. Die Ergebnisse wurden auf die Einschlusskriterien untersucht, unter anderem, ob sich die Studien auf Lehrkräfte auf allgemeinbildenden Schulen und nicht auf Lehramtsstudierende bezogen. Ebenfalls wurde analysiert, ob es sich bei der in den Studien betrachteten Technologie um Lernplattformen handelt, die zum digital-gestützten Üben genutzt werden können. Von 892 gefundenen Titeln, wurden 87 Abstracts und 37 Volltexte auf Passung analysiert.

Ergebnisse und ihre Bedeutung

Sechs Studien konnten als passend identifiziert werden. Drei Studien nutzen das TAM (Davis et al., 1989) und drei Studien das UTAUT Modell (Venkatesh et al., 2003) um die Technologieakzeptanz zu analysieren. Ebenfalls drei von sechs Studien wurden im Zeitraum der Covid-10 Pandemie durchgeführt. Aus den sechs eingeschlossenen Studien kann gefolgert werden, dass vor allem die Erfahrung im Umgang mit weniger komplexer Technologie für die Intention zur Nutzung von Technologie zum digital-gestützten Üben eine Rolle spielt, wie auch die Leistungs- und Aufwandserwartungen der Lehrkräfte, sowie das Vorhandensein von Support und der notwendigen digitalen Infrastruktur. Es konnte weiterhin ein Forschungsbedarf bezüglich der Betrachtung der Akzeptanz von Lehrkräften bezüglich der Nutzung von Technologie zu einem definierten und bestenfalls lernförderlichen Zweck identifiziert werden.

Sportspielvermittlung in der Grundschule: Synergien von TGfU und digital-gestützten Lehr-/Lernprozessen

Henrike Diekhoff, Steffen Greve, Jessica Süßenbach
Leuphana Universität Lüneburg, Deutschland

Theoretischer Hintergrund

Digitale Medien werden in der alltäglichen Lebenswelt der Kinder und Jugendlichen und damit auch in der Schule immer wichtiger. Das gilt auch für den Sportunterricht, da Medienpädagogik laut KMK Querschnittsaufgabe für alle Unterrichtsfächer ist (KMK, 2016). Die deutschsprachige Sportpädagogik nimmt sich aktuell dem Thema immer mehr an, bisher gibt es aber (nur) einige Praxisempfehlungen und wenig wissenschaftliche Begleitung, insbesondere für die Grundschule (Ausnahme: Greve et al., 2020). In internationalen Publikationen gibt es eine Vielzahl von Untersuchungen zu den Themen Gesundheit, kooperatives Lernen, Gamification und Wearable Technologies mit digitalen Medien. Dabei konzentrierte sich die Mehrzahl der Studien auf die Verbesserung der körperlichen Leistungsfähigkeit oder der körperlichen Aktivität und ging selten darüber hinaus (Jastrow et al., 2022). In dem beschriebenen Forschungsvorhaben wird die digital-gestützte Sportspielvermittlung für Spielanfänger:innen in der Grundschule untersucht. Grundlage ist der ganzheitliche Vermittlungsansatz des Teaching Games for Understanding (TGfU), der erstmals von Bunker und Thorpe als Gegensatz zum traditionellen Techniklernen vorgestellt wurde (Bunker & Thorpe, 1982).

Ziel ist es, die Schüler:innen zu flexiblen sowie kreativen Akteur:innen des eigenen Handelns und Denkens im Sportspiel auszubilden (Hopper, Butler & Storey, 2009). Die Untersuchung erfolgte im Hinblick auf die Frage, wie Schüler:innen den Einsatz digitaler Medien im Sportspielunterricht wahrnehmen und interpretieren. Hierbei wurden Bewegungsbildung, Allgemeinbildung und Medienbildung miteinander verschränkt und als gleichberechtigt interpretiert.

Methode

In zwei Zyklen der Datenerhebung wurde in sechs 4. Klassen über den Zeitraum von sechs Schulstunden eine Unterrichtseinheit zum Fußball nach der Vermittlungsmethode des TGfU durchgeführt. Während der Einheit wurden die Schüler:innen in Small-Sided-Games (3 vs. 3) mit dem Zielspiel konfrontiert, nahmen sich gegenseitig mit einer App auf und taggten Spielsituationen nach vorher festgelegten Kriterien. Die Sequenzen (ca. 8 Sekunden) wurden in den anschließenden Reflexionsphasen gemeinsam besprochen und auf dieser Grundlage initiierte die Lehrkraft für die Schüler:innen angemessene Spiel- und Übungsformen. Vor den Sportunterrichtsstunden fanden jeweils zwei Schulstunden Medienpädagogik statt, in denen die Kinder Zeit hatten die App auszuprobieren, Gesprächsregeln für den Umgang mit den Videoaufnahmen erarbeiteten und Themen wie das Recht am eigenen Bild besprochen wurden. Im Anschluss an die Einheit wurden 104 leitfadengestützte Interviews mit Schüler*innen geführt. Datenerhebung und -analyse erfolgen anhand der Grounded-Theory-Methodologie (Strauss & Corbin, 1996). Die Daten wurden offen und axial codiert und aus dem vorläufigen Kategoriensystem nach der ersten Erhebung wurde ein finales Kategoriensystem entwickelt.

Ergebnisse und ihre Bedeutung

Für die Schüler:innen war die Verwendung eines Tablets im Sportunterricht etwas Neues, dennoch waren die meisten Kinder mit der Bedienung der App schnell vertraut. Die Daten zeigen sehr ambivalente Sichtweisen auf die Nutzung der App (Greve et al., 2022). Die Schüler:innen gehen insgesamt sehr reflektiert mit dem Einsatz der App um. Das digitale Medium wird als Hilfsmittel im Lernprozess gedeutet und ist deutlich unwichtiger, als das Spielen selbst. Die Kinder berichten von vielen Versuchen, die in den Reflexionsphasen gemeinsam besprochenen taktischen Ideen im nachfolgenden Spiel umzusetzen. Interessant ist auch die Zweckentfremdung der Filmmöglichkeit: Die Schüler:innen nutzten diese Option oftmals als Videobeweis, um strittige Situationen schnell zu klären. Hier wird der Einfluss der mediatisierten Sportwelt am Beispiel des Wissens über den Videoschiedsrichter (VAR) deutlich. Die Ergebnisse zeigen somit eine überaus pragmatische Handhabung des digitalen Hilfsmittels durch die Schüler:innen, was im Vortrag mit Blick auf weitere Phänomene und Handlungsmuster diskutiert wird.

Feedback als Gelingensfaktor erfolgreichen Übens – Eine experimentelle Laborstudie zur Wahrnehmung und Wirksamkeit digitaler Lernplattformen

Larissa Altenburger, Michael Besser
Leuphana Universität Lüneburg, Deutschland

Theoretischer Hintergrund

Feedback wird nach Hattie und Timperley (2007) als „information provided by an agent [...] regarding aspects of one's performance or understanding" (S. 81) verstanden. Es gibt verschiedene Arten von

Feedback, die nach ihrer Komplexität geordnet werden können. Bei der Feedbackart „knowlege of result“ (Informationen über die Korrektheit einer Antwort) handelt es sich beispielsweise um eine wenig komplexe Art. Mehr Informationen (z. B. Hilfen) bietet elaboriertes Feedback (Shute, 2008).

Im Hinblick auf die Wirkung auf den Lernerfolg soll elaboriertes Feedback weniger komplexen Feedbackarten überlegen sein (Wisniewski et al., 2020) – wohingegen zu komplexes Feedback umgekehrte Effekte erzeugen kann (Duss, 2020). Diese Studienergebnisse lassen sich ebenfalls auf das Arbeiten mit digitalen Lernplattformen, die oftmals unterschiedlich komplexe Feedbackarten anbieten, übertragen (van der Kleij et al., 2015). Darüber wie Lernende dieses Feedback wahrnehmen, ist allerdings wenig bekannt. Neuere Studien legen jedoch die Notwendigkeit der unterstützenden Wahrnehmung von Feedback zur Leistungsentwicklung nahe. Diese Studien weisen auch auf eine positivere Wahrnehmung von elaboriertem Feedback im Vergleich zu weniger komplexem Feedback hin (Harks et al., 2013).

Fragestellung

Ausgehend von diesen theoretischen Vorüberlegungen befasst sich dieser Beitrag mit den folgenden Forschungsfragen:

(1) Wie nehmen Lernende verschiedene Feedbackarten einer Lernplattform wahr?

(2) Welchen Einfluss hat die Wahrnehmung des Feedbacks einer Lernplattform auf die Leistung?

Methode

Im Rahmen einer Laborstudie nahmen seit August 2021 insgesamt 147 (w = 67, d = 4) Siebtklässler:innen (Durchschnittsalter 12,5 Jahre) an einer Untersuchung zur Analyse zu Wahrnehmung und Wirksamkeit von Feedback aus digitalen Lernplattformen teil. Im Rahmen einer Vorerhebung wurden zunächst unter anderem die Bruchrechenleistung und das Geschlecht der Teilnehmenden zu Zwecken der parallelisierten, randomisierten Zuteilung zu zwei Versuchsgruppen (Vg1: knowledge of result; N = 74; Vg2: elaboriertes Feedback; N = 73) erhoben. Im Verlauf der Studie selbst arbeiteten beide Gruppen innerhalb einer entwickelten Intervention (Dauer: 30 Minuten) mit einer auf dem Bildungsmarkt etablierten, mathematischen Lernplattform an insgesamt 20 Aufgaben zur Bruchrechnung und erhielten – entsprechend der jeweiligen Bedingung – Feedback. Im Anschluss an die eigentliche Intervention wurden die Wahrnehmung des durch die Lernplattform angebotenen Feedbacks (5 Subskalen, Likert-Skalen mit 1 = Stimme überhaupt nicht zu bis 4 = Stimme völlig zu) sowie ein empirisch erprobter Bruchrechentest (nach Lenz et al., 2019) zur Erfassung der Bruchrechenleistung administriert.

Ergebnisse und ihre Bedeutung

Forschungsfrage 1 (Wahrnehmung):

Die Mittelwerte der fünf Wahrnehmungsskalen liegen für Vg1 zwischen 2.39 und 2.83 und für Vg2 zwischen 2.41 und 2.72. Diese Ergebnisse deuten auf eine mittelmäßige Wahrnehmung des Feedbacks hin. Außerdem weisen die Mittelwerte auf einen fehlenden signifikanten Unterschied der beiden Versuchsgruppen hinsichtlich ihrer Wahrnehmung hin. Ein t-Test bestätigt diese Vermutung.

Forschungsfrage 2 (Wahrnehmung auf Leistung):

Auch bezüglich der Bruchrechenleistung unterscheiden sich die Versuchsgruppen (MVg1 = 12.85, MVg2 = 13.22) nicht signifikant. Aber: Regressionsanalysen mit der wahrgenommenen Kompetenzunterstützung als unabhängige Variable und der Leistung der Schüler:innen im Bruchrechentest als abhängige Variable belegen einen Einfluss der Wahrnehmung elaborierten Feedbacks auf die Leistungsentwicklungen ($F(1,69) = 5.86$; $p = .02$; $\beta = .28$; $R2 = .078$). Vergleichbare Effekte für diejenigen Lernenden, die allein „knowledge of result" angeboten bekommen, liegen nicht vor.

Vorliegende Ergebnisse mögen im Kontext aufgezeigter theoretischer Überlegungen zunächst überraschen – zeigen sich doch keine direkten Unterschiede in der Wahrnehmung des Feedbacks sowie in den Lernfortschritten der Lernenden der verschiedenen Versuchsbedingungen. Dennoch sind diese Ergebnisse durchaus im Einklang mit bisheriger Forschung, dass Feedback zunächst als unterstützend wahrgenommen werden muss, bis dieses tatsächlich Lernfortschritte positiv beeinflussen kann (siehe Harks et al., 2013; Co2CA). Letztlich zeigt sich einmal mehr: Eine qualitativ hochwertige Implementation von Lernumgebungen kann – unabhängig von einer analogen und/oder digitalen Ausgestaltung – entscheidend für das Gelingen schulischer Lehr-Lern-Situationen sein.

«Wer hat's gemacht?» – Eine qualitative Studie zur Nutzung digitaler Lerndaten durch Grundschullehrkräfte

Alina Kristin Hase, Poldi Kuhl
Leuphana Universität Lüneburg, Deutschland

Theoretischer Hintergrund

Digitale Technologien bieten Lehrkräften Unterstützung bei der Individualisierung ihres Unterrichts. Beispielsweise bei der Bearbeitung von Übungsaufgaben gesammelte Fehleranalysen und Lernentwicklungsdaten haben das Potential, Lehrkräften zusätzliche Informationen über ihre Schüler:innen zur Verfügung zu stellen, die für pädagogische Entscheidungen genutzt werden können (Ebner & Ebner, 2018; Molenaar & Knoop-van Campen, 2017; Schaumburg, 2021). Bisher werden in Deutschland digitale Daten jedoch kaum zur Unterrichtsgestaltung genutzt (Blumenthal et al., 2021; Hase et al., 2022; Ifenthaler & Drachsler, 2020).

Im Kontext datenbasierter Entscheidungsfindung lässt sich die Nutzung von Daten für die Unterrichtsgestaltung in fünf Schritte unterteilen: (1) Problemidentifikation und Fragenformulierung, (2) Datennutzung, (3) Daten in Informationen umwandeln, (4) Informationen in Entscheidungen umwandeln, (5) Entscheidungsevaluation (Mandinach & Gummer, 2016).

Fragestellung

Ziel dieser Studie ist, zu überprüfen, ob sich der skizzierte Datennutzungskreislauf auf die Nutzung von Daten aus digitalen Lernplattformen übertragen lässt und beschreiben kann, wie Grundschullehrkräfte Daten aus digitalen Lernplattformen nutzen.

Methode

Zur Beantwortung der Forschungsfrage werden verschiedene Methoden herangezogen und kombiniert (Triangulation; Denzin, 1978). Im März 2022 wurden halbstrukturierte Interviews mit n=12 Grundschullehrinnen durchgeführt. Die interviewten Lehrkräfte hatten zuvor bereits an einer

Fragebogenstudie teilgenommen, in der die Nutzung von Daten quantitativ erfasst wurde (Hase et al., 2022). Der Interviewleitfaden der hier vorgestellten qualitativen Ergänzungsstudie umfasste Fragen zur Nutzung von Daten aus digitalen Lernplattformen, bspw. zur Ausgestaltung, zu Vor- und Nachteilen und benötigten Kompetenzen. Zusätzlich wurden in Einzelfallstudien mit n=2 der zuvor befragten Grundschullehrerinnen Beobachtungen ihres Unterrichts mit digitalen Lernplattformen sowie der Unterrichtsnachbereitung und Interviews durchgeführt. Die Interviewtranskripte und Beobachtungsdaten wurden einer inhaltlich strukturierenden qualitativen Inhaltsanalyse unterzogen (Kuckartz, 2018). Eine deduktive sowie induktive Kategorienbildung in MAXQDA führte zu thematischen Haupt- und Subkategorien.

Ergebnisse und ihre Bedeutung

Die von den Lehrkräften in den Interviews thematisierte Nutzung von Daten aus digitalen Lernplattformen konnte mit Hilfe des Datennutzungskreislaufs beschrieben und exemplarisch mittels der Einzelfallstudien in der Unterrichtspraxis belegt werden. Als Ausgangspunkt für die Datennutzung (Schritt 1) sahen die meisten interviewten Lehrkräfte die Frage, welche ihrer Schüler:innen die Übungsaufgaben erledigten und wie sie diese absolvierten. Nur vereinzelt gingen die Lehrkräfte mit einer spezifischeren Fragestellung an die Daten heran. Die beobachteten Lehrkräfte sahen ebenfalls hauptsächlich die Frage, ob und wie Aufgaben bewältigt wurden, als Grundlage für ihre 5- bis 20-minütige Betrachtung des Lehrkräfte-Dashboards der Lernplattform an. Alle interviewten Lehrkräfte beschrieben, welche Daten ihnen zur Verfügung standen und wie sie die Daten betrachteten (Schritt 2). Zwei Lehrkräfte blieben bei diesem Schritt stehen. Zehn Lehrkräfte berichteten von einer anschließenden Interpretation der Daten (Schritt 3). In der Beobachtung der Unterrichtsnachbereitung zeigte sich, dass das Nutzen der Daten mit der Interpretation der Daten zusammenfällt. Demnach überlegten sowohl zum Teil die interviewten als auch die beobachteten Lehrkräfte, was die Daten bedeuten, unter Rückbezügen zu ihrem Unterricht und ihren Schüler:innen. Wiederum sieben interviewte Lehrkräfte leiteten aus den gewonnenen Informationen schließlich Handlungsschritte ab (Schritt 4). Hierbei nannten Lehrkräfte beispielsweise die Anpassung nachfolgender Unterrichtsstunden oder Bereitstellung von individuellen Förder- und Forderaufgaben. Auch die beobachteten Lehrkräfte stellten ausgehend von den Daten Überlegungen für Unterrichtsveränderungen an. Eine Lehrkraft setzte die beschlossenen Maßnahmen, wie ein erneutes Aufgreifen der Unterrichtsthematik, teilweise um, wohingegen direkte Veränderungen des Unterrichts bei der zweiten Lehrkraft nicht beobachtet werden konnten. Der letzte Schritt des Datennutzungskreislaufes, die Entscheidungsevaluation, wurde von nur einer interviewten Lehrkraft thematisiert, sodass der Datennutzungskreislauf kaum vollständig durchlaufen wurde. Auch in den Einzelfallstudien war eine Evaluation nicht erkennbar.

Abschließend soll der Einblick in die Datennutzung von Grundschullehrkräften im Hinblick auf die Anforderungen in der Lehrkräfteaus- und -fortbildung diskutiert werden. Dabei steht im Fokus, ein Verständnis für eine vollumfänglichere Nutzung digitaler Daten aufzubauen. Somit von einer reinen «Wer hat’s gemacht?»-Intention zu dem Grundgedanken datenbasierter Entscheidungsfindung zurückzukehren, Daten für eine individualisierte Unterrichtsgestaltung zu nutzen.

B-18: Unterrichtserleben in Virtual Reality als Chance für die Lehrkräftebildung

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* T03 R03 D89

Unterrichtserleben in Virtual Reality als Chance für die Lehrkräftebildung

Chair(s): **Tosca Daltoè** (Hector-Institut für Empirische Bildungsforschung, Universität Tübingen & Institut für Bildungsanalysen Baden-Württemberg, Stuttgart), **Richard Göllner** (Hector-Institut für Empirische Bildungsforschung, Universität Tübingen)
*Diskutant*in(nen):* **Peter Gerjets** (Leibniz-Institut für Wissensmedien, Tübingen)

In der Lehrkräftebildung werden seit Jahren Unterrichtsvideos eingesetzt, um unterrichtliches Erleben und Handeln praxisnah darstellen und einüben zu können (Seidel & Thiel, 2017). In den letzten Jahren wurden klassische Unterrichtsvideos zunehmend durch neue Technologien, wie Virtual Reality (VR), ergänzt (Huang et al., 2021). Es stellt sich die Frage, welche Chancen, aber auch welche Herausforderungen der Einsatz von VR im Kontext Lehrkräftebildung mit sich bringt. Während VR-Umgebungen beispielweise ein erhöhtes Präsenzerleben ermöglichen, wird eine potenzielle mentale Überforderung dieser Technologie als Herausforderung diskutiert (Rupp et al., 2016). Das vorliegende Symposium stellt vier aktuelle Forschungsarbeiten zum Einsatz von VR in der Lehrkräftebildung dar:

Virtuelle 360°-Unterrichtssituationen – Immersive 360°-Unterrichtsvideos im Kontext der Professionellen Unterrichtswahrnehmung von angehenden Lehrpersonen im berufsbildenden Bereich

Im ersten Beitrag wird der Einsatz von 360°-Unterrichtsvideos in VR in der Lehrkräftebildung im berufsbildenden Bereich erprobt. In einer quasi-experimentellen Studie wurde untersucht, wie das VR-basierte Unterrichtserleben durch Studierende bewertet wird und ob sich diese Einschätzungen je nach unterrichtlicher Praxiserfahrung unterscheiden.

Mittendrin statt nur dabei – Wie unterscheiden sich das Unterrichtserleben und die Unterrichtsqualitätsratings von Lehramtsstudierenden am PC und in einer Virtual-Reality-Umgebung?

Der zweite Beitrag vergleicht die Unterrichtswahrnehmung von Lehramtsstudierenden in inszenierten 360°-Unterrichtsvideos zwischen einer PC- und einer VR-Bedingung. In einer experimentellen Studie wird untersucht, inwiefern sich das Unterrichtserleben sowie die aus der Unterrichtsbeobachtung resultierenden Unterrichtsqualitätsratings je nach Videoumgebung unterscheiden.

Maschinelles Lernen zur Analyse professioneller Unterrichtswahrnehmung im virtuellen Klassenzimmer

Der dritte Beitrag präsentiert einen Vergleich von aktiven Lehrkräften (Expert*innen) und Lehramtsstudierenden (Noviz*innen) hinsichtlich professioneller Unterrichtswahrnehmung in einem simulierten virtuellen Klassenzimmer. Mit Hilfe eines Machine-Learning-Ansatzes wurde analysiert, wie sich das Blick- und Bewegungsverhalten beim Umgang mit Unterrichtsstörungen je nach Expertise unterschied.

Videobasierte Unterrichtsreflexion in der Lehrkräfteausbildung: Virtual Reality und klassische Unterrichtsvideos im Vergleich

Der vierte Beitrag vergleicht simulierte VR-Unterrichtsvideos mit klassischen Unterrichtsvideos zur Förderung reflexionsbezogener Kompetenzen von Lehramtsstudierenden. In einer experimentellen Studie wurde untersucht, wie sich die reflexionsbezogene Selbstwirksamkeit und schriftliche

Unterrichtsreflexionen nach einer Videoreflexionsaufgabe zwischen einer VR-Gruppe und einer Standardvideo-Gruppe unterschied.

Mit diesen vier Beiträgen umfasst das Symposium Forschungsarbeiten zum unterrichtlichen Erleben und Handeln sowohl mit 360°-Unterrichtsvideos als auch in simulierten Unterrichtsumgebungen. Auf Basis der Beiträge soll übergeordnet auf einer kognitionspsychologischen Grundlage diskutiert werden, wie und wofür VR in der Lehrkräftebildung gewinnbringend eingesetzt werden kann.

Beiträge des Symposiums

Virtuelle 360°-Unterrichtssituationen – Immersive 360°-Unterrichtsvideos im Kontext der Professionellen Unterrichtswahrnehmung von angehenden Lehrpersonen im berufsbildenden Bereich

Katharina Kunz, Bernd Zinn
Abteilung Berufspädagogik mit Schwerpunkt Technikdidaktik, Institut für Erziehungswissenschaft, Universität Stuttgart

Theoretischer Hintergrund

Videobasierte Fallarbeit findet vielfältige Einsatzmöglichkeiten zur Theorie-Praxis-Verzahnung, um der Komplexität von Unterricht bereits in der ersten Lehrpersonenbildungsphase gerecht zu werden und um eine Erprobung von handlungsbezogenem Wissen in unterrichtlichen Kontexten zu ermöglichen (Gaudin & Chaliès, 2015). Vorteile liegen u. a. in der hohen Authentizität und der standardisierten Wiederholbarkeit von ganzen Unterrichtssituationen durch Videos (Syring et al. 2015). Neue Technologien, wie 360°-Videos in Virtual Reality (VR) können im Vergleich zu traditionellen Videos authentischere Perspektiven, Situierung und Immersion ermöglichen und stellen eine interessante Erweiterung des beforschten Bereichs der Fallarbeit durch den Einsatz einer neuen Educational Technology dar (Kunz & Zinn 2022; Rosendahl & Wagner, 2021).

Das Vorhaben fokussiert die Erprobung von virtuellen unterrichtlichen Situationen (über 360°-Videos), die als Erfahrungsräume Praxisfelder im Lehramtsstudium ergänzen können und einen praxis- und wissenschaftsorientierten Kompetenzerwerb im Rahmen hochschulischer Lehr- und Lernarrangements ermöglichen. Die grundsätzliche Eignung von 360°-Unterrichtsvideos in VR zur Entwicklung von Interventionen zur Förderung der Professionellen Unterrichtswahrnehmung bei heterogenen Lerngruppen im berufsbildenden Bereich soll anhand der folgenden Forschungsfragen untersucht werden:

1. Wie bewerten Studierende der beruflichen Bildung die Technologieakzeptanz, Immersion und Präsenzerleben von 360°-Videos in VR im Kontext ihrer eigenen Professionalisierung?

2. Wie nützlich bewerten Studierende der beruflichen Bildung 360°-Videos in VR hinsichtlich der eigenen Professionalisierung und welche explorativen Entwicklungsperspektiven verbinden sich damit?

Methode

In einem quasi-experimentellen Setting wurden quantitative und qualitative Zugänge genutzt, um die Verhaltensmuster der Studierenden zu Technologieakzeptanz, Immersion und Präsenzerleben durch standardisierte Beobachtung zu erfassen und um kategoriale Daten zur Nützlichkeit und Entwicklungsperspektiven zu 360°-Videos in VR zu generieren. Erfahrene Lehrpersonen entwickelten

Unterrichtsszenarien, welche in einem Unterrichtsraum auf Basis von Videoskripten mit Darsteller*innen mit einer 360°-Kamera aufgezeichnet wurden. Die Unterrichtsszenen zeigen eine kleine Lerngruppe einer Förderklasse einer beruflichen Schule. Die Testpersonen durchliefen alle 360°-Videos in VR: 1. Erleben der Unterrichtsszenen aus der Egoperspektive über die VR-Brille Occulus Go mit der Aufgabe, die Rolle der Lehrperson in den Unterrichtsszenen reaktiv einzunehmen. 2. Die Testpersonen reflektierten die eigene Reaktion verbal. 3. Erhebung quantitativer Daten zu den Konstrukten Immersion, Präsenzerleben und Technologieakzeptanz sowie qualitativer Daten zu Entwicklungsperspektiven und Nützlichkeit hinsichtlich der Professionalisierung. Es nahmen N = 57 Testpersonen teil, die zum Zeitpunkt der Testung Berufs- oder Technikpädagogik studierten, wobei 21 Personen (36,8%) bereits praktische Unterrichtserfahrung gesammelt haben.

Ergebnisse

Die Ergebnisse der Skalen Technologieakzeptanz, Immersion und Präsenzerleben weisen eine durchschnittlich hohe Bewertung durch die Testpersonen auf, was eine insgesamt positive Wahrnehmung der 360°-Videos in VR durch die Studierenden impliziert. Die hohe Immersion ist interessant, da so Emotionen sowie das Präsenzerleben beeinflusst werden können und damit möglicherweise eine lernförderliche Wirkung erzielt werden kann (Ramalho & Chambel, 2013). Insgesamt können hohe Immersion und Präsenzerleben durch ihre Bedeutung für die Situierung als ein entscheidender Faktor für einen erfolgreichen Lernprozess eingeordnet werden, da davon ausgegangen wird, dass der Lernprozess durch hohe Immersion und ein hohes wahrgenommenes Präsenzerleben positiv beeinflusst wird (Hebbel-Seeger 2018). Die Ergebnisse der Subgruppenvergleiche mittels t-Tests (der Studierenden mit und ohne Praxiserfahrung) zeigen, dass Studierende der Berufs- und Technikpädagogik ohne unterrichtliche Praxiserfahrung eine höhere Nützlichkeit, Nutzungsintention und Präsenzerleben angeben (Kunz & Zinn 2022).

Die qualitativen Ergebnisse umfassen insgesamt N = 148 Kodierungen in den Frageaspekten „subjektive Bewertung zur Nützlichkeit für die eigene Professionalisierung (100 Codes)“ und „Entwicklungsperspektiven zur Technologienutzung (48 Codes)“. Die Auswertung erfolgte anhand der qualitativen Inhaltsanalyse nach Mayring (2010; Cohen’s Kappa k = 0,74). Die Ergebnisse belegen die Wahrnehmung der virtuellen Unterrichtsszenarien als geschützten Lern- und Erprobungsraum, da das eigene Handeln nicht einer unmittelbaren Beobachtung durch Dozierende ausgesetzt ist. Der situative Realitätsbezug und immersive Charakter der Interaktion wird als förderlich bewertet. Das geschützte Erproben von problemhaltigen Unterrichtssituationen wird fast durchweg als nützlich für die eigene Professionalisierung genannt, insbesondere durch Studierende ohne Unterrichtserfahrung (Kunz & Zinn 2022).

Mittendrin statt nur dabei – Wie unterscheiden sich das Unterrichtserleben und die Unterrichtsqualitätsratings von Lehramtsstudierenden am PC und in einer Virtual-Reality-Umgebung?

Tosca Daltoè[1], Evelin Ruth-Herbein[2], Ann-Kathrin Jaekel[3], Benjamin Fauth[1], Ulrich Trautwein[3], Richard Göllner[3]
[1]Hector-Institut für Empirische Bildungsforschung, Universität Tübingen & Institut für Bildungsanalysen Baden-Württemberg, Stuttgart, [2]Institut für Bildungsanalysen Baden-Württemberg, Stuttgart, [3]Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

Theoretischer Hintergrund

Als Repräsentationen der Praxis werden Unterrichtsvideos in der Lehrkräftebildung und Forschung seit vielen Jahren eingesetzt (Gaudin & Charliès, 2015; Seidel & Thiel, 2017). Unterrichtsvideos bilden das komplexe unterrichtliche Geschehen authentisch ab und bieten den Vorteil, einzelne Unterrichtssequenzen gezielt fokussieren und mehrfach betrachten zu können (Syring et al., 2015). Bisherige Unterrichtsvideoforschung hat gezeigt, dass sich Unterrichtsvideos dazu eignen, Kompetenzen von angehenden Lehrkräften zu fördern, wie beispielsweise die professionelle Unterrichtswahrnehmung (Weber et al., 2018) oder die Feedbackkompetenz (Prilop et al., 2020).

Neben ihrem großen Potenzial für die Lehrkräftebildung haben klassische Unterrichtsvideos durchaus auch Limitationen. Studien aus der Unterrichtsforschung zeigen beispielsweise, dass die Kameraperspektive, aus der der Unterricht gefilmt wird, die Wahrnehmung des videographierten Unterrichts beeinflusst (Cortina et al., 2018; Paulicke et al., 2019). Eine Möglichkeit, dieser Herausforderung entgegenzutreten, ist der Einsatz von 360°-Videos. 360°-Unterrichtsvideos ermöglichen eine Unterrichtsbeobachtung, bei der die Perspektive auf den Unterricht vom Standpunkt der Kamera aus frei wählbar ist (Balzaretti et al., 2019). 360°-Unterrichtsvideos können zudem in einer immersiven Virtual-Reality-Umgebung (VR-Umgebung) betrachtet werden, was in einer reduzierten psychologischen Distanz zur Videoumgebung resultieren kann (z. B. Breves & Schramm, 2021).

Erste empirische Studien zum Einsatz von 360°-Unterrichtsvideos am PC weisen auf ein höheres Präsenzerleben (Gold & Windscheid, 2020) sowie eine erleichterte Wahrnehmung zentraler Unterrichtsevents (Kosko et al., 2021) im Vergleich zu Standard-Unterrichtsvideos hin. Der Einsatz von 360°-Unterrichtsvideos, insbesondere in VR-Umgebungen, ist allerdings noch nicht hinreichend erforscht. Es wird angenommen, dass die Betrachtung von 360°-Unterrichtsvideos in VR-Umgebungen zu einem höheren Immersionserleben führt. Dabei stellt sich die Frage, ob ein immersives Unterrichtserleben in VR für das Unterrichtserleben und resultierende Unterrichtsqualitätsratings förderlich oder eher hinderlich ist. Die VR könnte durch die hohe Immersion entweder zu einem authentischeren Unterrichtserleben und einer adäquateren Einschätzung der Unterrichtsqualität oder aber zu Überforderung und einer weniger adäquaten Unterrichtseinschätzung führen. Ziel der präsentierten Studie ist es, anhand von inszenierten Unterrichtsvideos zur kontrastierenden Darstellung konkreter Unterrichtsqualitätsmerkmale des Unterrichtsfeedbackbogen Tiefenstrukturen (UFB; Fauth et al., 2021) zu untersuchen, inwiefern sich das Unterrichtserleben in Unterrichtsvideos und die resultierenden Unterrichtsqualitätsratings zwischen einer Standard-PC-Bedingung und einer 360°-VR-Bedingung unterscheidet.

Methode & Ausblick auf Ergebnisse

Für den Vergleich von Unterrichtserleben und Unterrichtsqualitätsratings wurden im Rahmen des Videoprojekts zum UFB auf der Grundlage von Unterrichtsskripts inszenierte 360°-Unterrichtsvideos an Schulen in Baden-Württemberg produziert. Die inszenierten Unterrichtsvideos stellen Positiv- und Negativbeispiele von Unterrichtsqualitätsmerkmalen der drei Basisdimensionen der Unterrichtsqualität dar, die mit dem UFB erfassbar sind. Anhand der inszenierten Unterrichtsvideos untersucht die präsentierte experimentelle Studie, ob sich Unterschiede im Unterrichtserleben und in resultierenden

Unterrichtsqualitätsratings von Lehramtsstudierenden zwischen einer Standard-PC-Bedingung und einer 360°-VR-Bedingung zeigen. Hinsichtlich des Unterrichtserlebens werden Mittelwertsunterschiede innerhalb der Studierenden zwischen der PC- und der VR-Bedingung in der wahrgenommenen Authentizität der Unterrichtssituation, im Immersions- und Präsenzerleben sowie in der wahrgenommenen mentalen Beanspruchung untersucht. Hinsichtlich der wahrgenommenen Unterrichtsqualität werden Mittelwertsunterschiede zwischen den Studierenden zwischen der PC- und der VR-Bedingung in den Unterrichtsqualitätsratings mit dem UFB (Fauth et al., 2021) betrachtet.

Zum Zeitpunkt der Einreichung liegen die inszenierten Unterrichtsvideos, jedoch noch keine Daten der experimentellen Studie vor. Die Studie wird im Wintersemester 2022/2023 mit Master-Lehramtsstudierenden (Zielgröße von N = 100) an der Universität Tübingen durchgeführt. Zum Zeitpunkt der Tagung werden die dann hochaktuellen Daten und Ergebnisse präsentiert.

Relevanz

Die Befunde der präsentierten Studie liefern wichtige Erkenntnisse zum Einsatz von Unterrichtsvideos in VR-Umgebungen in Lehrkräftebildung und Unterrichtsqualitätsforschung. Die Ergebnisse können zur Beantwortung der Frage beitragen, ob VR-Umgebungen bisherige Ansätze zur Repräsentation von Unterrichtspraxis in Lehrkräftebildung und Unterrichtsqualitätsforschung gewinnbringend erweitern können.

Maschinelles Lernen zur Analyse professioneller Unterrichtswahrnehmung im virtuellen Klassenzimmer

Philipp Stark[1], Hong Gao[2], Patricia Goldberg[1], Efe Bozkir[2], Enkelejda Kasneci[3], Richard Göllner[1]
[1]Hector-Institut für Empirische Bildungsforschung, Universität Tübingen, [2]Department of Computer Science, Human-Computer Interaction, Universität Tübingen, [3]TUM School of Social Sciences and Technology, Technische Universität München

Theoretischer Hintergrund und Fragestellung

Virtuelle Realitäten (VR) ermöglichen Verhaltensmessungen von Personen in einer kontrollierten und standardisierten Umgebung. Gleichzeitig schafft eine virtuelle Umgebung eine realitätsnahe Darstellung alltäglicher Situationen, in welcher Personen eine hohe Verhaltensauthentizität zeigen (vgl. Jensen & Konradsen, 2018) und Fertigkeiten praxisnah trainieren können (Blömeke et al., 2015; Grossman et al., 2009). Entsprechend bietet VR auch für Lehrkraftprofessionalisierung ein ausgesprochen hohes Potential (z. B. Huang et al., 2022; Sabers et al., 1991; Spencer et al., 2019).

Aktuelle VR-Systeme erlauben es zudem, unterschiedliche Verhaltensdaten, wie etwa Blickverhalten und Bewegungen (z. B. durch die Nutzung eines Controllers), während des Erlebens einer VR aufzuzeichnen. Die resultierenden Daten können zur Analyse der professionellen Unterrichtswahrnehmung genutzt werden, um zu erkennen, wie Lehrkräfte wesentliche Merkmale von Unterrichtssituationen wahrnehmen und interpretieren (Seidel & Stürmer, 2014). Beispielsweise lässt sich untersuchen, wie Lehrkräfte mit Störsituationen im virtuellen Klassenzimmer umgehen (Carrasco, 2011; Minarikova et al., 2021). Allerdings stellt die Verarbeitung dieser reichhaltigen Daten neue Herausforderungen an die Datenanalyse. Blick- und Bewegungsdaten zeichnen sich besonders durch ihren mehrdimensionalen Charakter aus (zeit-, personen-, und messgrößenabhängig). Es wird eine Vielzahl an Variablen gemessen, welche oft in einem nicht-linearen Zusammenhang stehen und komplex interagieren. In der Folge werden immer häufiger Methoden des maschinellen Lernens (ML)

eingesetzt (Hilbert et al., 2021). ML bietet die Möglichkeit, diese Daten zu strukturieren und Verhaltensmuster zu identifizieren.

In unserer Studie untersuchten wir in einem virtuellen Klassenzimmer das Blick- und Bewegungsverhalten von angehenden Lehrkräften (Novizen) und aktiven Lehrkräften (Experten) im Hinblick auf ihren Umgang mit Unterrichtsstörungen. Konkret wurde mithilfe von ML untersucht, welche Unterschiede sich im Blick- und Bewegungsverhalten zwischen Experten und Novizen in Bezug auf die Wahrnehmung und den Umgang mit Störungen im virtuellen Klassenzimmer ergaben.

Methode

Es wurden die Blick- und Bewegungsdaten von N = 42 Probanden (25 angehende Lehrkräfte, 17 aktive Lehrkräfte) analysiert, welche in einem virtuellen Klassenzimmer eine 10-minütige Präsentation vor einer virtuellen Klasse mit 18 realistisch animierten Schülerinnen und Schülern hielten. Die simulierten Schülerinnen und Schüler zeigten unterschiedliche Formen von Störverhalten (aktiv, passiv). Erkannte Störungen konnten von den Lehrkräften mittels des Controllers beendet werden. Die Unterrichtssituation war in 20 unterschiedlich komplexe Störintervalle eingeteilt, für welche wir 36 Merkmale des Blick- und Bewegungsverhaltens berechneten. Mithilfe eines Random-Forest ML-Verfahrens wurden zwei Kompetenzstufen (Experten und Novizen) klassifiziert. Die für die Klassifikation relevanten Blick- und Bewe-gungsdaten konnten mithilfe eines erklärbaren ML-Ansatzes (Shapley) identifiziert werden (Adadi & Berrada, 2018).

Ergebnisse und ihre Bedeutung

Der verwendete ML-Ansatz war in der Lage, Experten und Novizen in 78% der Fälle an der Blick- und Bewegungsdaten korrekt zu klassifizieren. Im Erkennen der störenden Schülerinnen und Schüler zeigte sich, dass Novizen die Störenden weniger oft stoppten, weniger oft anschauten (Anzahl Fixationen auf Störende) und kürzer mit ihrem Blick auf ihnen verweilten (Fixationsdauer auf Störende). Im Gegensatz dazu fixierten Novizen generell länger und öfter auf die virtuellen Schüler als Experten (Fixationen auf virtuelle Schülerinnen und Schüler all-gemein). Zudem zeigten Novizen einen höheren mittleren Pupillendurchmesser, welcher positiv mit einer höheren mentalen Anstrengung assoziiert ist (z. B. Krejtz et al., 2018).

Insgesamt zeigten die Studienergebnisse, dass Experten ein gezielteres Blickverhalten aufwiesen. In der herausfordernden Unterrichtssituation waren Experten besser in der Lage, störende Schülerinnen und Schüler während des Vortrags zu identifizieren (vgl. van Driel et al., 2021). Daraus lässt sich ebenfalls ableiten, dass unsere VR-Umgebung dazu geeignet ist, Kompetenzunterschiede von Lehrkräften in einer komplexen Unterrichtssituation abzubilden. Im nächsten Schritt soll die in der vorliegenden Studie erprobte VR-Umgebung auch in der Praxis eingesetzt werden. Die Nutzung von ML-basierten Analysemodellen ermöglicht die Verarbeitung komplexer Datenstrukturen, welche in der Zukunft die Entwicklung automatischer Systeme zu Erfassung der professionellen Unterrichtswahrnehmung von Lehrkräften möglich macht.

Videobasierte Unterrichtsreflexion in der Lehrkräfteausbildung: Virtual Reality und klassische Unterrichtsvideos im Vergleich

Eric Richter, Isabell Hußner, Yizhen Huang, Dirk Richter, Rebecca Lazarides
Universität Potsdam, Department Erziehungswissenschaft

Theoretischer Hintergrund

Die Fähigkeit und Bereitschaft zur Reflexion eigenen Unterrichts stellt eine wichtige Voraussetzung für das professionelle Handeln von Lehrkräften dar (Imhof & Schlag, 2018). Reflexion wird dabei zumeist als eine Form des Denkens verstanden, die sowohl kognitive als auch metakognitive Prozesse umfasst (Clarà, 2015). Diese Reflexionsprozesse können verschiedene Schritte beinhalten und sich auf unterschiedliche Gegenstände beziehen (Wyss, 2018). Hinsichtlich möglicher Reflexionsschritte unterscheiden Kleinknecht und Gröschner (2016) die Beschreibung einer pädagogischen Situation, die Bewertung und Begründung des/der Handelnden und die Formulierung von Handlungsalternativen. Hinsichtlich möglicher Reflexionsgegenstände postulieren Lohse-Bossenz et al. (2019), dass sich Unterrichtsreflexionen auf die Handlungen der Lehrkraft oder der Schüler*innen sowie auf das Lernumfeld beziehen können.

Reflexionsprozesse ermöglichen es Lehrkräften, ihr unterrichtliches und pädagogisches Handeln zu betrachten, zu hinterfragen und anzupassen (Gläser-Zikuda et al. 2018). Auf diese Weise können Reflexionsprozesse dazu beitragen, tradierte Denk- und Handlungsmuster aufzubrechen. Ein Ziel der Lehrkräfteausbildung besteht daher darin, Lehramtsstudierenden die Fähigkeit und Bereitschaft zu vermitteln, ihre Unterrichtserfahrungen systematisch zu reflektieren. Eine bedeutsame reflexionsbezogene Kompetenz ist in diesem Zusammenhang die reflexionsbezogene Selbstwirksamkeit (Lohse-Bossenz et al., 2019).

Eine Möglichkeit zur Förderung reflexionsbezogener Kompetenzen stellt der Einsatz von Unterrichtsvideos dar (Hamel & Viau-Guay, 2019). Sie dienen als “springboards for analysis and discussions about teaching and learning” (Borko et al., 2011, S. 184) und ermöglichen durch Realitätsnähe, Informationsfülle und Anschaulichkeit, gezielt die Komplexität von Unterrichtssituationen zu beurteilen (Krammer & Reusser 2005). Videobasierte Unterrichtsreflexionen setzen jedoch das Vorhandensein authentischer Unterrichtsvideos voraus. Insbesondere eigene Unterrichtsvideos der Studierenden stehen dabei häufig nicht zur Verfügung oder können nur unter hohem Einsatz von Ressourcen aufgezeichnet werden. Vor diesem Hintergrund kann der Einsatz von Virtual Reality (VR) Lehramtsstudierenden und Dozierenden die Möglichkeit bieten, Unterricht in einer virtuellen Umgebung zu üben, authentische Unterrichtserfahrungen in Klassenzimmern zu sammeln und diese letztlich zu videographieren (Wiepke et al., 2019).

Ziel der vorliegenden Studie ist es zu untersuchen, inwiefern die Anwendung solcher VR-Unterrichtsvideos zu Reflexionszwecken Lehramtsstudierende in der Entwicklung ihrer reflexionsbezogenen Selbstwirksamkeit unterstützt. Zudem soll überprüft werden, ob sich schriftliche Unterrichtsreflexionen in Abhängigkeit des zugrundeliegenden Videos (VR-Unterrichtsvideo vs. reales Unterrichtsvideo) hinsichtlich der zugrundeliegenden Reflexionsprozesse (Reflexionsschritte, Reflexionsgegenstand) unterscheiden.

Methode

Die vorliegende Studie im experimentellen Prä-Post-Design untersucht die Veränderung der selbstberichteten reflexionsbezogenen Selbstwirksamkeit (t1α = 0,85; t2α = 0,82; Lohse-Bossenz et

al., 2019) von N = 69 Lehramtsstudierenden (n = 46 VR-Gruppe, n = 23 Kontrollgruppe) mittels Varianzanalysen mit Messwiederholung. Zusätzlich wurden n = 68 schriftliche Unterrichtsreflexionen mithilfe qualitativer Inhaltsanalyse (Mayring & Frenzl, 2019) in Bezug auf die Reflexionsschritte und den Reflexionsgegenstand ausgewertet, die entweder auf Grundlage des eigenen VR-Videos (VR-Gruppe) oder anhand eines Unterrichtsvideos einer fremden Lehrkraft (Kontrollgruppe) verfasst wurden.

Ergebnisse

Die reflexionsbezogene Selbstwirksamkeit der Lehramtsstudierenden stieg nach der Videoreflexionsaufgabe in beiden Gruppen an (MVR $\Delta t1, t2 = 0{,}12$; MKG $\Delta t1, t2 = 0{,}07$). Ein Post-hoc-Test ergab, dass dieser Zuwachs für die VR-Gruppe von t1 zu t2 bei mittlerer Effektgröße signifikant ausfiel. Für die Kontrollgruppe war dies nicht der Fall.

Die Ergebnisse der qualitativen Inhaltsanalyse verweisen darauf, dass sich die schriftlichen Unterrichtsreflexionen der Studierenden in der VR-Gruppe und in der Kontrollgruppe hinsichtlich der Reflexionsschritte (Beschreiben, Begründen/Bewerten, Alternativen formulieren) nicht voneinander unterschieden. Unterrichtsreflexionen, die von Studierenden der VR-Gruppe verfasst wurden, fokussierten jedoch häufiger die lehrende Person (Lehramtsstudierende selbst) und seltener die Lernumgebung im Vergleich zu Unterrichtsreflexionen, die von Studierenden der Kontrollgruppe verfasst wurden.

Relevanz

Die Befunde verdeutlichen das Potenzial von VR-Anwendungen für die Lehrkräfteausbildung. Sie zeigen, dass VR-Unterrichtsvideos geeignete Lerngelegenheiten für den Erwerb reflexionsbezogener Kompetenzen darstellen und ähnliche Reflexionsprozesse ermöglichen wie reale Unterrichtsvideos. Insgesamt sind die Bedingungen, unter denen sich Selbstwirksamkeitserwartungen in der Lehrerkräfteausbildung verändern, bisher jedoch nur unzureichend empirisch untersucht worden (Clark & Newberry, 2019).

B-19: Emotionen und Emotionsregulation von Lehramtsstudierenden und deren Bedeutung für Professionalisierungsprozesse

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* R11 T00 D03

Emotionen und Emotionsregulation von Lehramtsstudierenden und deren Bedeutung für Professionalisierungsprozesse

Chair(s): **Irina Rosa Kumschick** (Pädagogische Hochschule Luzern), **Gerda Hagenauer** (Universität Salzburg, Austria)
*Diskutant*in(nen):* **Manfred Holodynski** (Universität Münster)

Emotionen steuern das Unterrichtsverhalten von Lehrpersonen, deren Wohlbefinden und Professionalisierungsprozesse (Frenzel, 2014; Hagenauer & Hascher, 2022). Diese Zusammenhänge gelten ebenso für Lehramtsstudierende, wenn sie—meist im Zuge von Schulpraktika—ihre ersten Unterrichterfahrungen machen. Schulpraktika werden von Lehramtsstudierenden hoch emotional erlebt (Anttila et al., 2016; Waber et al., 2021). Dies ist unter anderem mit der hohen Relevanz dieser Erfahrungen für die Überprüfung der Eignung für die Profession zu begründen. Gleichzeitig können die Studierenden i.d.R. auf eine geringe Erfahrung beim Unterrichten zurückgreifen, wodurch auch Fragen der Kontrollierbarkeit der Situation aufgeworfen werden. Es liegt folglich ein Setting vor, das aufgrund der Wertzuschreibung und Kontrolleinschätzung (Kontroll-Wert-Theorie; Pekrun, 2006) hohes emotionales Potential aufweist.

Aus der Erziehungswissenschaft und Psychologie werden aktuelle Arbeiten vorgestellt, die sich mit dem emotionalen Erleben und der Emotionsregulation (Gross, 2015) von Lehramtsstudierenden beim Unterrichten auseinandersetzen. Ziel dieses Symposiums ist es, emotionale und kognitive Prozesse, wie beispielsweise den Umgang mit Emotionen, die unterrichtsbezogene Selbstwirksamkeitsüberzeugung und den Einbezug der Perspektive von Schüler*innen als Forschungsinhalte zu untersuchen und deren Relevanz im Hinblick auf Professionalisierungsprozesse im Lehramtsstudium zu diskutieren.

In Beitrag 1 werden Angst, Ärger und Freude beim Unterrichten untersucht, und es wird geklärt, inwiefern diese durch das klassenbezogene Verhalten der Schüler*innen und die individuellen Kontroll- und Wertüberzeugungen der Studierenden erklärt werden können. Für die Angst erweist sich weniger das Verhalten der Schüler*innen als bedeutsam, sondern die individuellen Kontroll- und Wertappraisals der Studierenden.

Dieser Befund wird in den Beiträgen 2 und 3 aufgegriffen, wo der Zusammenhang zwischen den Emotionen und der Selbstwirksamkeitsüberzeugung beim Unterrichten näher beleuchtet wird. Es wird angenommen, dass Emotionen eine wesentliche Informationsfunktion für die Selbstwirksamkeitsüberzeugung aufweisen (Bandura, 1997) und die Selbstwirksamkeitsüberzeugung ihrerseits wiederum das Emotionserleben mitbestimmt. Die Ergebnisse belegen diesen Zusammenhang überwiegend, wobei distinkte Emotionen auch unterschiedlich stark mit den verschiedenen Facetten der Selbstwirksamkeitsüberzeugung zusammenhängen (Beitrag 2). Zudem zeigen die Ergebnisse von Beitrag 3, dass insbesondere gering selbstwirksame Studierende ihre Unterrichtserfahrungen—auch die als positiv bewerteten—emotional negativer erleben. Eine mögliche Ursache könnte in einer ungünstigen Situationsbewertung liegen.

Daher sollten Studierende Unterrichtssituationen adäquat bewerten. Dieser Aspekt wird in Beitrag 4 aufgegriffen. Auf der Basis eines Experiments kann gezeigt werden, dass handlungsbezogene

Kognitionen mit einer adaptiveren Emotionsregulation einhergehen, während personenbezogene Kognitionen maladaptive Regulationsstrategien begünstigen. Auch die Übernahme der Schüler*innenperspektive steht positiv mit der Emotionsregulation in Verbindung.

Die Beiträge werden abschließend im Hinblick auf ihre Bedeutung für das Forschungsfeld und die Ausgestaltung des Lehramtsstudiums, insbesondere der schulpraktischen Phasen, zur Anregung optimaler Professionalisierungsprozesse diskutiert.

Beiträge des Symposiums

Emotionen von Studierenden im Unterrichtspraktikum

Eva Becker[1], Tina Hascher[2]
[1]Universität Zürich, [2]Universität Bern

Theoretischer Hintergrund.

Aus Sicht der Lehramtsstudierenden zählen Unterrichtspraktika zu den wichtigsten Lernorten ihrer Ausbildung und werden überwiegend positiv bewertet (Hascher, 2006). Die (temporäre) Verantwortung für eine Klasse zu übernehmen und die Komplexität des Unterrichtens direkt zu erleben, kann jedoch auch eine emotionale Herausforderung sein (Sutton & Wheatley, 2003). Zwar zeigten einzelne quantitative Studien, dass vorwiegend positive Emotionen im Praktikum erlebt oder erwartet werden (de Zordo et al., 2019; Timoštšuk et al., 2016), doch deuten qualitative Befunde darauf hin, dass negative Emotionen, insbesondere Angst, im Unterrichtspraktikum durchaus präsent und handlungsleitend sind (Hastings, 2004; Timoštšuk & Ugaste, 2012) und die retrospektive Beurteilung des Praktikums prägen bzw. verzerren können. Aus diesem Grund ist der Einsatz ergänzender situativer Erhebungsmethoden als auch zusätzlicher Datenquellen (z.B. Befragung der Schülerinnen und Schüler) indiziert.

Fragestellung.

Auf der Grundlage eines theoretischen Modells (Frenzel, 2014), untersucht die vorliegende Studie die Antezedenzien der diskreten Emotionen Freude, Angst und Ärger von Studierenden im Praktikum. Dabei werden sowohl distale Variablen des Lern- und Leistungsverhalten der unterrichteten Klasse als auch subjektive Bewertungsprozesse (in Bezug auf Kontroll- und Wertüberzeugungen) der Studierenden sowie personen- und kontextbezogene Variablen hinzugezogen, um die folgenden Forschungsfragen zu beantworten: 1) Welche Emotionen erleben Studierende in den durchgeführten Lektionen? 2) Wie stark hängen die Emotionen mit dem situativen Verhalten der Praktikumsklasse sowie mit den Kontroll- und Wertüberzeugungen in der jeweiligen Unterrichtslektion zusammen? 3) Welche Vorerfahrungen der Studierenden und Kontextvariablen des Praktikums stehen mit den Kontroll- und Wertüberzeugungen der Studierenden und ihren Emotionen im Zusammenhang?

Methode.

Die vorliegende Studie kombiniert eine Fragebogenuntersuchung (nach dem Praktikum) der Lehramtsstudierenden (N = 173, MAlter = 26.98, 68.2% weiblich) mit Tagebucherhebungen (Befragungen im Anschluss an die Lektionen) der Studierenden und ihren Praktikumsklassen (N = 3616 Schüler*innen, Jahrgang 2007-2010, MAlter=14.84, 52.8% weiblich). Es wurden Daten in Deutschland und der Schweiz während eines (mindestens) dreiwöchigen Unterrichtspraktikums im

Fach Deutsch erhoben. Die Studierenden führten jeweils eine Unterrichtseinheit (sechs Lektionen) zum Thema „schriftliches Argumentieren“ durch und dokumentierten nach drei dieser Lektionen ihre unterrichtsbezogenen Emotionen (Einzelitems basierend auf den TES; Frenzel et al., 2016) sowie ihre Kontrollüberzeugungen (6 Items, α = .82, z.B. «In dieser Lektion hatte ich das Gefühl, alles unter Kontrolle zu haben») und Wertüberzeugungen (7 Items, α = .70, z.B. «In dieser Lektion war es mir wichtig, gut zu unterrichten»). Gleichzeitig berichteten die Schüler*innen über Unterrichtsstörungen und ihr Interesse während derselben Unterrichtsstunden (ICC2 = .82 / .71). Im abschließenden Fragebogen nach der Unterrichtseinheit berichteten die Studierenden demographische und studienbezogene Daten, ihre bisherige Unterrichtserfahrung sowie die erhaltene Unterstützung im Praktikum (z.B. Anzahl Unterrichtsbesprechungen).

Ergebnisse.

Die Studierenden berichteten in der Mehrzahl der untersuchten Unterrichtsstunden (80.2%) von intensiver Freude (Wert ≥ 4 auf einer fünfstufigen Skala), aber auch gelegentlich von intensiven Ängsten (14.2%) oder Ärger (7.8%). Insbesondere das Auftreten intensiver Ängste/Nervosität ist damit etwa doppelt so häufig wie bei erfahrenen Lehrpersonen (vgl. Frenzel, 2014). In mehrebenenanalytischen Regressionsmodellen zeigten die Emotionen Freude und Ärger im Unterricht signifikante Zusammenhänge mit den berichteten Unterrichtsstörungen sowie dem Interesse der Klasse. Das Erleben von Angst zeigte sich unabhängig vom erfassten Verhalten der Klasse. Der Einbezug der Kontroll- und Wertüberzeugungen führte zu einer systematischen Steigerung der Varianzaufklärung (auf 20-25%) und scheint insbesondere für das Erleben von Unterrichtsängsten relevant. Kontroll- und Wertüberzeugungen zeigten zudem signifikante Zusammenhänge zu den Vorerfahrungen der Studierenden (weitere Unterrichtserfahrungen und Vorkenntnisse im Themengebiet der Unterrichtseinheit). Die Ergebnisse weisen darauf hin, dass die Emotionen von Studierenden im Praktikum deutlich weniger stark durch das tatsächliche Unterrichtsgeschehen (beurteilt durch die Klasse) beeinflusst werden, sondern vielmehr von ihren eigen Wert- und Kontrollüberzeugungen abhängen.

Im Vortrag soll die Rolle angeleiteter Reflexionsphasen (z.B. mit Peers oder der Praktikumslehrperson) zur Einordnung der Emotionserfahrungen und Regulationsmöglichkeiten der Studierenden diskutiert werden.

Freude, Ärger und Angst im Schulpraktikum: Wie hängen Emotionen mit der Selbstwirksamkeit von Lehramtsstudierenden zusammen?

Andreas Bach, Gerda Hagenauer
Universität Salzburg

Theoretischer Hintergrund.

Lehramtsstudierende erleben in Schulpraktika zahlreiche Emotionen (Hascher & Hagenauer, 2016). Aufgrund unterschiedlicher Situationsbewertungen im Sinne eines appraisaltheoretischen Ansatzes (Ellsworth & Scherer, 2003) ist davon auszugehen, dass diese interindividuell variieren, wobei positive Emotionen wie Freude zu dominieren scheinen (Bach, 2022; Bach & Hagenauer, 2021, 2022) – aber auch Ärger und Angst können auftreten (Morton et al., 1997; Porsch & Gollub, 2018). In der Forschung wurden Emotionen als Antezedenz und Folge von Selbstwirksamkeit theoretisiert und

empirisch untersucht (Bandura, 1997; Chen, 2019; Pekrun, 2006). Der Vortrag nimmt explizit Bezug zu Banduras Selbstwirksamkeitstheorie (Bandura, 1997). Emotionale Zustände bilden in diesem theoretischen Rahmen insbesondere eine von vier Quellen für die Entstehung von Selbstwirksamkeit. Da in der Forschung zu Schulpraktika bislang mehrheitlich negative Emotionen als eindimensionaler Faktor untersucht wurden (Morris et al., 2017; Pfitzner-Eden et al., 2014), ist empirisch weitgehend ungeklärt, welche Effekte unterschiedliche—sowohl positive als auch negative—Emotionen auf die Selbstwirksamkeit von Studierenden in verschiedenen Domänen haben.

Fragestellung.

Dieses Forschungsdesiderat aufgreifend untersuchte die Studie, wie die von Studierenden in einem Unterrichtspraktikum erlebten Emotionen Freude, Ärger und Angst mit Selbstwirksamkeitsüberzeugungen bezogen auf Instruktionsstrategien, Classroom Management und Schüler*innen-Engagement zusammenhängen.

Methode.

Befragt wurden 449 Lehramtsstudierende einer österreichischen Universität (66.6% weiblich, Alter: M = 23.38, SD = 4.17), die mehrheitlich (72.6%) im fünften Bachelor-Semester Lehramt Sekundarstufe studierten und die in der Regel zwei semesterbegleitende Unterrichtspraktika in den von ihnen studierten Unterrichtsfächern absolvierten. Zur Erfassung der Emotionen wurde ein adaptierter Fragebogen von Frenzel et al. (2016) genutzt (Freude: $\alpha = .80$; Ärger: $\alpha = .77$; Angst: $\alpha = .78$; jeweils vier Items). Die Selbstwirksamkeitsüberzeugungen wurden mit einer Skala von Pfitzner-Eden et al. (2014) erfasst (Instruktionsstrategien: $\alpha = .66$; Classroom Management: $\alpha = .87$; Schüler*innen-Engagement: $\alpha = .74$; jeweils vier Items). Konfirmatorische Faktorenanalysen sprachen für die dreifaktorielle Struktur beider Skalen. Berechnet wurde ein latentes Regressionsmodell mit Freude, Ärger und Angst als unabhängige Variablen und den drei Dimensionen von Selbstwirksamkeit als abhängige Variablen. Die statistischen Analysen wurden mit dem Programm Mplus 8.5 (Muthén & Muthén, 1998–2017) durchgeführt. Die Parameterschätzungen basierten auf dem Maximum-Likelihood-Robust(MLR)-Verfahren. Fehlende Werte wurden mittels Full-Information-Maximum-Likelihood (FIML)-Verfahren berücksichtigt.

Ergebnisse und ihre Bedeutung.

Freude war für alle untersuchten Domänen von Selbstwirksamkeit förderlich. Der stärkste Effekt zeigte sich für Schüler*innen-Engagement ($\beta = .42$, $p < .001$), gefolgt von Instruktionsstrategien ($\beta = .33$, $p < .001$) und Classroom Management ($\beta = .19$, $p = .001$). D.h., je mehr Freude Studierende im Praktikum erlebten, desto stärker schätzten sie ihre Fähigkeiten ein, Aufgaben in diesen unterrichtlichen Verhaltensbereichen durchzuführen. Ärger stand in einem negativen Zusammenhang mit der Selbstwirksamkeit hinsichtlich des Schüler*innen-Engagements ($\beta = –.21$, $p = .006$). Angst wiederum war negativ mit der klassenführungs- ($\beta = –.47$, $p < .001$) und instruktionsbezogenen ($\beta = –.37$, $p < .001$) Selbstwirksamkeit assoziiert. Je mehr Ärger und Angst Studierende im Praktikum erlebten, desto geringer schätzten sie ihre Fähigkeiten ein, Schüler*innen zu motivieren, mit Störungen im Unterricht umzugehen und Instruktionsstrategien anzuwenden. In Übereinstimmung mit den theoretischen Annahmen (Bandura, 1997) und bisherigen empirischen Befunden (Bach, 2022;

Bach & Hagenauer, 2021, 2022; Pfitzner-Eden, 2016), bestätigt die Studie, dass sowohl negativ als auch positiv erlebte Emotionen im Praktikum mit der Selbstwirksamkeit von Studierenden zusammenhängen.

Die Befunde erweitern den Forschungsstand, indem empirische Erkenntnisse in Bezug auf spezifische Emotionen und Domänen von Selbstwirksamkeit vorliegen. Die Studienergebnisse und Implikationen für die Lehrer*innenbildung zum Umgang mit Emotionen im Praktikum werden im Vortrag vorgestellt und diskutiert.

Emotionsbezogene Wortwahl in Unterrichtsreflexionen und ihre Zusammenhänge zu unterrichtsbezogenen Selbstwirksamkeitserwartungen von Lehramtsstudierenden

Anja Henke[1], Andrea Westphal[2], Isabell Hußner[1], Rebecca Lazarides[1]
[1]Universität Potsdam, [2]Universität Greifswald

Theoretischer Hintergrund.

Die Kontroll-Wert-Theorie von Pekrun (2006) beschreibt leistungsbezogene Emotionen als proximale Prädiktoren von Kompetenzerwerb. Häufig wurden die Annahmen der Kontroll-Wert-Theorie im Kontext Schule untersucht. Es kann allerdings auch angenommen werden, dass Emotionen in anderen Bildungskontexten bedeutsam für Lernprozesse sind. So verweisen Befunde zu Lehramtsstudierenden auf positive Zusammenhänge zwischen positiven professionsbezogenen Emotionen und ihrer professionsbezogenen Selbstwirksamkeitserwartung (Brígido et al., 2013). Die schriftliche Reflexion von Unterrichtssituationen ist eine in der Lehramtsausbildung häufig genutzte Möglichkeit, um erlebte Emotionen zu beschreiben (Pollmeier et al., 2021). Allerdings wurde bislang nur wenig untersucht, inwiefern sich Emotionen in schriftlichen Reflexionen von Unterrichtserfahrungen äußern. Experimentalpsychologische und klinische Studien zeigen, dass sich das emotionale Erleben von Individuen in ihrer Wortwahl widerspiegelt, wenn sie persönlich relevante Situationen reflektieren. Eine offene Frage ist, inwiefern die Identifikation des Emotionsgehalts schriftlicher Unterrichtsreflexionen dazu beitragen kann, die Unterrichtspraxis von Lehramtsstudierenden zu verbessern und ihr Kompetenzerleben zu fördern.

Fragestellungen.

Die vorliegende Studie untersucht, inwiefern in schriftlichen Reflexionen von unterschiedlich wahrgenommenen Unterrichtssituationen (positiv beziehungsweise herausfordernd), auch unterschiedliche emotionsbezogene Wörter (positive beziehungsweise negative) gewählt werden. Zudem wird untersucht, inwiefern die emotionsbezogene Wortwahl mit unterrichtsbezogenen Selbstwirksamkeitserwartungen von Studierenden zu Beginn und zum Ende eines Semesters zusammenhängen.

Methode.

Genutzt wurden Daten von N = 146 Lehramtsstudierenden (53.6% weiblich, 46.4% männlich; MAlter = 23.92, SD = 5.05), die an einem praxisorientierten Seminar mit Microteaching-Elementen teilnahmen. Die Studierenden reflektierten schriftlich zwei Unterrichtssituationen – eine als positiv wahrgenommene und eine als herausfordernd wahrgenommene Situation. Unterrichtsbezogene Selbstwirksamkeitserwartungen wurden orientiert an Pfitzner-Eden et al. (2014) zu Beginn und am Ende des Semesters erhoben (Subskalen: Schülermotivation, Instruktionsstrategien,

Klassenmanagement). Mit der Software Linguistic Inquiry and Word Count (LIWC; Meier et al., 2018) wurde die Verwendung von positiven und negativen emotionsbezogenen Wörtern in den N = 239 schriftlichen Reflexionen bestimmt. Zur Auswertung wurden mit der Statistiksoftware Mplus 8.7 drei separate Pfadmodellierung für die drei Subskalen der Selbstwirksamkeitserwartungen vorgenommen.

Ergebnisse.

Die Analysen zeigten, dass Studierende, die in als herausfordernd wahrgenommenen Situationen eine negativere emotionsbezogene Wortwahl nutzen, auch in als positiv wahrgenommenen Situationen eine negativere emotionsbezogene Wortwahl nutzen (in den drei Pfadmodellierungen: $\beta \geq 0.22$; $SE \leq 0.09$; $p \leq .011$). Zudem zeigte sich, dass Lehramtsstudierende, die zur Mitte des Semesters in schriftlichen Reflexionen von positiv wahrgenommenen Unterrichtssituationen mehr negative emotionsbezogene Wörter nutzen, auch geringere Selbstwirksamkeitserwartungen (SW) zum Ende des Semesters berichten (SWSchülermotivation: $\beta = -0.23$; $SE = 0.09$; $p = .011$; SWInstruktionsstrategien: $\beta = -0.24$; $SE = 0.08$; $p = .002$; SWKlassenmanagement: $\beta = -0.17$; $SE = 0.07$; $p = .020$), auch wenn für die Selbstwirksamkeitserwartungen zu Beginn des Semesters kontrolliert wird. Überraschend ist, dass sich keine signifikanten Zusammenhänge zwischen positiven emotionsbezogenen Wörtern in schriftlichen Reflexionen – als herausfordernd wahrgenommene und als positiv wahrgenommene Situationen – und den Selbstwirksamkeitserwartungen ergaben.

Relevanz.

Die vorliegende Studie verdeutlicht, dass es für die Lernprozesse der Lehramtsstudierenden lohnenswert ist, ihre Emotionen in Praxisreflexionen zu berücksichtigen. Dabei tendieren Lehramtsstudierende mit geringen Selbstwirksamkeitserwartungen dazu, negative emotionsbezogene Worte zu wählen, wenn sie als positiv wahrgenommene Unterrichtserfahrungen reflektieren. Möglicherweise können diese Personen eigene Erfolgserfahrungen («mastery experiences») nicht als solche erkennen. Praktische Implikationen für die Lehrkräftebildung betreffen die pädagogische Diagnostik und Identifikation jener Lehramtsstudierenden, die positive Praxiserfahrungen als negativ wahrnehmen. Diese Gruppen Studierender sollten systematische personalisierte Unterstützung erhalten, die ihnen die akkurate Wahrnehmung von Erfolgserfahrungen ermöglicht.

Wie sich eine handlungsbasierte versus personenbezogene Kognition auf die Emotionsregulation während Unterrichtsstörungen auswirkt

Irina Rosa Kumschick[1], Loredana Torchetti[2], Martina Lang[1], Annette Tettenborn[1]
[1]Pädagogische Hochschule Luzern, [2]Pädagogische Hochschule Bern

Theoretischer Hintergrund.

Unterrichten bedeutet, multiple Anforderungen in einer interaktiven Situation unter Handlungsdruck erfolgreich und adäquat zu bewältigen. Insbesondere der Umgang mit störendem Schüler*innenverhalten im Klassenzimmer stellt für Lehrpersonen eine der am schwersten zu bewältigenden Berufsanforderungen dar (Admiraal et al., 2000; Gallagher, 2009). Die affektiven und verhaltensbezogenen Reaktionen von Lehrpersonen auf das Fehlverhalten ihrer Schüler*innen hängen stark davon ab, wie sie Störungen im Unterricht erklären (Reyna & Weiner, 2001). Gleichzeitig hat die Unfähigkeit von Lehrpersonen, mit Unterrichtsstörungen erfolgreich umzugehen, nicht nur negative Folgen für sie selbst (Chan, 2006), sondern auch für die Lernenden (Sutherland et

al., 2008) und die Beziehung zwischen der Lehrperson und ihren Schüler*innen (Hamre & Pianta, 2005).

Fragestellungen.

In diesem Experiment wurde untersucht, wie Lehrpersonen Unterrichtsstörungen kognitiv erklären und emotional regulieren. Es wurde die Hypothese untersucht, ob Lehrpersonen eine funktionalere Emotionsregulation (Gross & John, 2003; Hochschild, 1983) zeigen, wenn ihre Kognition handlungsbasiert ist und wenn sie die Perspektive der Schüler*innen auf die Unterrichtsstörung kennen.

Methode.

An der Studie nahmen N = 167 Lehramtsstudierende im Masterstudiengang für Lehrpersonen der Sekundarstufe I teil. Allen angehenden Lehrpersonen wurde ein Video mit einer stark gestörten Unterrichtssituation gezeigt. Im Vergleich zu den Teilnehmenden der Kontrollgruppe, wurde den Teilnehmenden der Experimentalgruppe anschließend ein zusätzliches Video präsentiert, in dem die/der störende Schüler*in ihr/sein Verhalten kommentierte und erklärte. Im Anschluss füllten alle Teilnehmenden der Studie Fragebögen zu ihrer kognitiven Verarbeitung der Unterrichtsstörung aus. Dabei ging es um handlungsbasierte Kognitionen (z. B. Nachdenken über den eigenen Unterrichtsstil) oder personenbezogene Gedanken (z. B. Nachdenken über eigene Inkompetenz). Zudem wurden Fragen zur Emotionsregulation (Cognitive Change, Attentional Deployment, Deep Acting und Venting) gestellt.

Ergebnisse.

Es hat sich gezeigt, dass Studierende mit handlungsbasierten Kognitionen signifikant häufiger funktionale Emotionsregulationsstrategien nutzten (Cognitive Change: $\beta = .46$, Attentional Deployment: $\beta = .33$, Deep Acting: $\beta = .44$). Gleichzeitig wendeten Teilnehmende mit einer pessimistischen personenbezogenen Kognition häufiger eine dysfunktionale Regulation mit Venting ($\beta = .35$) und signifikant weniger häufig Deep Acting ($\beta = -.19$) zur Regulation ihrer Emotion an. Ferner zeigte sich, dass Teilnehmende, welche die Störenden- Perspektive kannten, häufiger Cognitive Change als Emotionsregulationsstrategie nutzten als Teilnehmende der Kontrollgruppe ($\beta = .38$). Darüber hinaus zeigte sich ein Trend dahingehend, dass Lehramtsstudierende, welche die Störenden-Perspektive kannten—auch wenn sie eine pessimistische personenbezogene Kognition hatten—seltener die dysfunktionale Emotionsregulation Venting benutzten ($p = .061$).

Schlussfolgerung.

Die Ergebnisse sprechen dafür, dass es sinnvoll ist, Studierende explizit in handlungsbasierten Kognitionen zu trainieren, da diese mit funktionaleren Emotionsregulationsstrategien—im Sinne von Cognitive Change, Attentional Deployment und Deep Acting—verbunden sind. Zudem hat die vorliegende Studie gezeigt, dass die Perspektive des/der in der Störsituation involvierten Schülers/Schülerin ein strukturierendes, integratives, regulierendes Moment für den kognitiven Umgang mit Emotionen darstellt, das nicht nur dazu führt, dass eine dysfunktionale Regulation

negativer Emotionen vermindert wird, sondern auch dazu, dass bedeutend häufiger Cognitive Change genutzt wird.

B-20: Was tun? Perspektiven für eine Unterrichtsqualitätsforschung der Zukunft

Zeit: Dienstag, 28.02.2023: 15:15 - 17:00 · *Ort:* R11 T00 D05

Was tun? Perspektiven für eine Unterrichtsqualitätsforschung der Zukunft

Chair(s): **Lukas Begrich** (DIPF), **Anna-Katharina Praetorius** (Universität Zürich, Switzerland), **Mareike Kunter** (DIPF), **Sandy Taut** (Bayerisches Landesamt für Schule)
*Diskutant*in(nen):* **Marten Clausen** (Universität Duisburg-Essen)

Die deutschsprachige Unterrichtsqualitätsforschung hat sich in der Vergangenheit sowohl in empirischen als auch theoretischen Arbeiten mit der Frage beschäftigt, was Unterrichtsqualität kennzeichnet. Sie hat Modelle hervorgebracht, die im deutschsprachigen Raum breit rezipiert werden und als eine gemeinsame Grundlage für die Erforschung der Merkmale und Bedingungen erfolgreichen Unterrichts dienen. Die konzeptionellen und methodischen Herausforderungen, die sich insbesondere durch die Komplexität des Untersuchungsgegenstandes Unterricht ergeben, werden durch die Unterrichtsqualitätsforschung im Sinne einer stetigen Weiterentwicklung des eigenen theoretischen wie methodischen Repertoires adressiert (vgl. z.B. Praetorius, Grünkorn & Klieme, 2020). Gleichzeitig erscheint fraglich, inwieweit ihre Erkenntnisse auf schulischen Unterricht anwendbar sind, der im Zuge gesellschaftlicher Veränderungen in deutlichem Ausmaß im Wandel begriffen ist. Eine fortschreitende Digitalisierung, zunehmende Heterogenität, Klimawandel, Globalisierung, Zulauf für antidemokratische Bewegungen etc. verändern sowohl die Kompetenzen, die Kinder und Jugendliche für ein erfolgreiches Leben brauchen, als auch die Art, wie Kinder und Jugendliche lernen, und in welchen Kontexten sie lernen (Koch, 2022; OECD, 2019; Zweck et al., 2015). Um auch zukünftig Erkenntnisse gewinnen zu können, die zum Gelingen von Unterricht beitragen, scheint ein Perspektivwechsel erforderlich, in dem die Wandlungsprozesse des Unterrichts antizipiert und daraus notwendige Innovationen für die Unterrichtsqualitätsforschung abgeleitet werden. Als einen ersten Schritt in diese Richtung hat das Leibniz-Netzwerk Unterrichtsforschung einen Diskurs zwischen Expert*innen des Feldes über Entwicklungsmöglichkeiten und -notwendigkeiten der Unterrichtsqualitätsforschung angeregt. In dem Symposium sollen vier Arbeitsgruppen einen Einblick in diesen Diskurs geben.

Jasmin Decristan und Kolleg*innen setzen sich in ihrem Beitrag damit auseinander, wie sich Unterricht durch die steigende Heterogenität der Schüler*innenschaft und damit verknüpften veränderten Rollenverständnissen auf Seiten der Lehrkräfte und Schüler*innen verändert und wie dem methodisch adäquat begegnet werden kann.
Marc Kleinknecht und Kolleg*innen diskutieren in ihrem Beitrag – ebenfalls vor dem Hintergrund einer zunehmenden Heterogenität der Lernenden und der daraus resultierenden Wichtigkeit adaptiven Unterrichtens – die zentrale Bedeutung von Diagnostik im Unterricht und die Notwendigkeit, zukünftig die unterschiedlichen diesbezüglichen theoretischen und empirischen Ansätze zusammenzuführen.

Christian Herrmann und Kolleg*innen zeigen in ihrem Beitrag Möglichkeiten auf, wie ein fachübergreifendes Rahmenmodell von Unterrichtsqualität mit fachspezifischen Ausdifferenzierungen in einem iterativen Prozess mit der Unterrichtspraxis entwickelt und implementiert werden kann.

Richard Göllner und Kolleg*innen beleuchten in ihrem Beitrag notwendige methodische Innovationen für die Unterrichtsqualitätsforschung.

Unterricht im Kontext sich auflösender Strukturen

Jasmin Decristan[1], Doris Holzberger[2], Eckhard Klieme[3], Susanne Kuger[4], Christine Pauli[5], Katrin Rakoczy[6]
[1]Bergische Universität Wuppertal, [2]Technische Universität München, [3]DIPF, [4]DJI, [5]Universität Freiburg (CH), [6]Justus Liebig Universität Gießen

Bereits seit mehr als einem Jahrzehnt ist der Umgang mit einer zunehmend heterogenen Schüler*innenschaft ein prominentes Thema von Bildungspraxis und Unterrichtsforschung (vgl. Dumont, 2019; Klieme & Warwas, 2011). So ist individuelle Förderung inzwischen nicht nur in allen Schulgesetzen verankert, sondern die Schulen selbst setzen in ihrem Unterricht auch verstärkt Maßnahmen der Differenzierung und Individualisierung um. Dabei werden den Akteur*innen mehr Freiheitsgrade und Flexibilität eingeräumt, sowohl in den strukturell-organisatorischen Unterrichtsabläufen (z.B. durch Freiarbeitszeiten und weniger starke Taktungen statt fixer Stundenpläne), in den Sozialformen (z.B. durch flexible Lerngruppen und freie Wahl der Lernpartner*innen statt einem festen Klassensetting) und bei den Inhalten (z.B. durch flexible Ausgestaltung des Curriculums und Verknüpfung verschiedener Fachinhalte statt strikt nach Fächern getrenntem Unterricht). Damit einher gehen ein verändertes Rollenverständnis sowohl der Lehrkräfte (im Sinne einer Supervisorin und Moderatorin des Lernens; vgl. Pauli & Reusser, 2000) als auch der Schüler*innen, bei denen verstärkt auf eigenverantwortliches („choice and voice of learning“, z.B. Mötteli et al., 2022) und selbstreguliertes Lernen gesetzt wird. Ebenso werden nicht nur fachliche Leistungen der Schüler*innen in den Blick genommen, sondern Bildungs- und Entwicklungsziele ganzheitlicher betrachtet. Erste Einblicke in die Schulen auf Basis empirischer Studien, wie Ada*Q (Adaptivität und Unterrichtsqualität im individualisierten Unterricht; Decristan, Dumont & Fauth, 2017) und PerLen (Personalisiertes Lernen in heterogenen Lerngruppen; Stebler, Pauli & Reusser, 2018) legen nahe, dass dieser aus den Schulen heraus stattgefundene bottom-up-Prozess in einem Blumenstrauß an mehr oder weniger kreativen Konzepten mündet. Jedoch liegt aus der Unterrichtsforschung hierzu erstaunlich wenig Empirie vor.

An diesem Desiderat setzt der vorliegende Beitrag an und stellt die Diskussionen aus dem Leibniz-Netzwerk Unterrichtsforschung zu zentralen Aufgaben und Herausforderung der zukünftigen Unterrichtsforschung vor. Folgende übergreifende Fragen sind hierfür leitend:

1. Was charakterisiert Unterricht im Kontext sich auflösender Strukturen?

2. Wie lässt sich individualisierter Unterricht methodisch erfassen?

3. Welche empirischen Wirkungen lassen sich hierzu aufzeigen?

Um diese zentralen Fragen zu adressieren, gilt es zunächst zu explizieren, wie diese „neue Adaptivität“ mit ihren ganz unterschiedlichen Begrifflichkeiten (z.B. Personalisiertes Lernen, Neue Lernkultur) an bisherige Begriffe und Konzepte der Erziehungswissenschaft und Pädagogischen Psychologie anknüpft (z.B. Dumont, 2019) und mit welchen theoretischen Grundannahmen diese jeweils in Beziehung gesetzt werden können. Ebenso bedarf es vertiefter Überlegungen, was qualitätsvollen Unterricht unter den veränderten Strukturen charakterisiert. Es lässt sich erwarten, dass die generischen Dimensionen der kognitiven Aktivierung, konstruktiven Unterstützung und Klassenführung weiterhin zentrale Qualitätsmerkmale konstituieren, dass diese jedoch in ihren

Konzeptualisierungen und Operationalisierungen stärker auf die individualisierten Angebotsstrukturen angepasst werden müssen. Diese theoretisch-konzeptionellen Überlegungen leiten zu zentralen methodischen Herausforderungen der zukünftigen Unterrichtsforschung über. Es bedarf einer Ergänzung, Anpassung und teils auch innovativen Neu-Entwicklung von Erhebungsinstrumenten, um die Flexibilität, die unterrichtlichen Interaktionen und die Individualisierung des Lernens zu erfassen. Multiple Kamerasysteme (Troll et al., 2022) und mobile Tischkameras sind hierfür erste Umsetzungsbeispiele. Für valide Aussagen sollten quantifizierende Methoden mit einer dichteren Beschreibung der unterrichtlichen Interaktionen durch qualitative Verfahren (z.B. Breidenstein & Rademacher, 2017) angereichert werden. Fragebögen für Schüler*innen müssen um multiple Ziele ergänzt werden, sodass zusätzlich zu fachlichen Leistungen auch sozio-emotionale und motivationale Ziele, aber auch Persönlichkeitsentwicklung, Demokratiebildung und Kompetenzen im Umgang mit gesellschaftlichen und ökologischen Veränderungen (Digitalisierung, Vielfalt, Klima) einbezogen werden. Zudem gilt es das soziale Gefüge (z.B. über Verfahren der Netzwerkanalyse; Troll et al., 2022; Zander et al., 2017) eingehender in den Blick zu nehmen, auch um bspw. das Spannungsfeld aus Individualisierung und Gemeinschaft näher zu beleuchten. Um differenzielle Wirkungen abhängig von den Lernvoraussetzungen zu erfassen, kann auf die ursprüngliche Idee der Aptitude-Treatment-Interaktionen (vgl. Corno & Snow, 1986) zurückgegriffen und diese weiterentwickelt werden. Zusammen genommen soll der Beitrag neue Impulse und Diskussionen für die zukünftige empirische Forschung zu Unterricht in einer zunehmend heterogenen Schüler*innenschaft ermöglichen.

Diagnostik und guter Unterricht

Marc Kleinknecht[1], Benjamin Fauth[2], Esther Brunner[3], Anke Lindmeier[4], Alexander Naumann[5]
[1]Leuphana Universität Lüneburg, [2]benjamin.Institut für Bildungsanalysen Baden-Württemberg, [3]Pädagogische Hochschule Thorgau, [4]Friedrich-Schiller-Universität Jena, [5]DIPF

Eine vielfältig konstatierte Heterogenität in den Lernvoraussetzungen von Schülerinnen und Schülern verlangt von Lehrkräften ein adaptives Unterrichten, bei dem Diagnostik eine wesentliche Rolle spielt. Diagnostik umfasst analytisches Beobachten von Lernaktivitäten ebenso wie eine Lernstatusdiagnostik. Diese hohe Bedeutung der Diagnostik für die Gestaltung adaptiven, lernwirksamen Unterrichts erfordert aus einer wissenschaftlichen Perspektive das Zusammenführen von theoretischen Ansätzen und empirischen Analysen, die bislang zumeist in relativ disparaten Forschungsfeldern diskutiert wurden. Dabei gilt es drei Forschungsfelder, die sich mit der Rolle von Diagnostik im Unterricht befasst haben, in den Blick zu nehmen.

Erstens beschäftigt sich die pädagogisch-psychologische Diagnostik, mit der Frage, wie sich psychometrisch orientierte Verfahren im Unterricht und für den Unterricht nutzen lassen (Hesse & Latzko, 2017). So bieten etwa Instrumente der Lernverlaufsdiagnostik die Möglichkeit, individuelle Lernverläufe grafisch abzubilden und bei auffälligen Entwicklungen individuell intervenieren zu können. Zweitens ist diagnostisches Handeln als Teil des Unterrichtshandeln Gegenstand der empirischen Unterrichtsqualitätsforschung. Es wird überall dort mitgedacht, wo die Konzeptualisierung von Unterrichtsqualität mehr oder weniger explizit beinhaltet, dass das Unterrichtsangebot passend sein muss zu den individuellen Voraussetzungen der Lernenden, etwa um kognitive Aktivierung oder individuelle Unterstützung zu ermöglichen (Sliwka, Klopsch & Dumont, 2022). So erfordert etwa ein lernwirksames Feedback diagnostische Informationen zum Vorwissen und möglichen Wissenslücken

einer Schülerin bzw. eines Schülers. Drittens sind Fragen der Nutzung diagnostischer Informationen für die Unterrichtsgestaltung auch immer schon selbstverständlicher Bestandteil fachdidaktischer Forschungen (Leuders et al., 2022). Diese Ansätze sind unverzichtbar, weil sie für einzelne Unterrichtsinhalte fachspezifisches deklaratives, prozedurales und konzeptuelles Wissen sowie seine Bedeutung für den Lernprozess spezifizieren können, was wiederum notwendige Grundlage für die Entwicklung diagnostischer Instrumente ist.

Zum Teil überschneiden sich Themen zwischen diesen drei Forschungsfeldern, oftmals stehen die jeweiligen Überlegungen jedoch noch unverbunden nebeneinander, obgleich zum Teil ähnliche Problemstellungen adressiert werden. Es ist daher eine Aufgabe für eine interdisziplinär denkende Unterrichtsforschung zu eruieren, wie sich die oben genannten Forschungsstränge zusammenführen ließen, um theoretische Überlegungen und empirische Befunde für die konkrete Praxis der Lehrkräftebildung und der Unterrichtsentwicklung fruchtbar zu machen. In diesem Zusammenhang scheint es lohnenswert, von konkreten Unterrichtspraktiken auszugehen, in denen diagnostische Vorgehensweisen relevant werden, etwa wenn Lehrkräfte Unterrichtsgespräche führen oder Lernende beim Erschließen von Fachtexten anleiten (Kleinknecht, Krammer & Seel, in Druck) oder wenn sie fachliche Fehlkonzepte in Antworten von Schülerinnen und Schüler erkennen oder solche antizipierend bei der Unterrichtsplanung berücksichtigen. Von einer partiellen Synthese der oben genannten theoretischen Ansätze könnten Vorschläge ausgehen, wie solche Praktiken weiterentwickelt und verbessert werden könnten.

Entwicklung eines fachübergreifenden Rahmenmodells von Unterrichtsqualität mit fachspezifischen Ausdifferenzierungen

Christian Herrmann[1], Anna-Katharina Praetorius[2], Christiane Bertram[3], Anke Lindmeier[4], Mirjam Steffensky[5], Patrick Schreyer[6], Iris Winkler[4], Stefan Keller[1]
[1]Pädagogische Hochschule Zürich, [2]Universität Zürich, [3]Universität Konstanz, [4]Friedrich-Schiller-Universität Jena, [5]Universität Hamburg, [6]Universität Kassel

Theoretischer Hintergrund

Die Unterrichtsqualitätsforschung entwickelte verschiedene Konzeptualisierungen zur Beschreibung von Unterrichtsqualität, welche als weitestgehend stufen- und fachübergreifend (generisch) begriffen wurden (im Überblick Praetorius & Charalambous, 2018). Dabei stellt sich die Frage, ob diese generischen Modelle für die verschiedenen Schulfächer Gültigkeit besitzen, da sich die Lernziele, -gegenstände und -prozesse zwischen den Fächern teils stark unterscheiden. In aktuellen Übersichtsarbeiten wurde deutlich, dass die verschiedenen (Sub-)Dimensionen der Unterrichtsqualität fachspezifisch ausdifferenziert werden müssen, um die Unterrichtsprozesse in den verschiedenen Fächern adäquat beschreiben zu können (u.a. Praetorius et al., 2020). Damit rückt eine fachdidaktische Unterrichtsforschung in den Fokus, welche ausgehend vom jeweiligen Gegenstand des Faches (z. B. Bewegung, Experimente, Kommunikation) spezifische Lehr- und Lernprozesse beschreiben und konkrete (Sub-)Dimensionen guten Fachunterrichts ableiten. Was beispielsweise eine gelingende kognitive Aktivierung ausmacht, wird in und zwischen den Fächern sehr kontrovers diskutiert. So werden in Abhängigkeit der didaktischen und lerntheoretischen Fundierung der Fächer unterschiedliche Kognitionen als bedeutsam erachtet, um fachbezogene Lernprozesse bei Schüler*innen zu unterstützen. Diese Heterogenität von fachspezifischen und fachübergreifenden

Modellen der Unterrichtsqualität stellt eine Hürde für den Diskurs über die Fachdisziplinen hinweg dar, erschwert die Aus- und Weiterbildung der Lehrpersonen und schafft für die Unterrichtspraxis keine klaren Orientierungslinien. Damit stellt sich die Arbeitsgruppe der Frage, ob sich ein Rahmenmodell qualitätsvollen Fachunterrichts erarbeiten lässt.

Methode

In mehreren aufeinander aufbauenden Austauschtreffen der fach- und disziplinübergreifenden Arbeitsgruppe grenzen in einem ersten Schritt Fachdidaktiker*innen aus unterschiedlichen Fächern (u.a. Englisch, Mathematik, Deutsch, Sport, Chemie, Geschichte) und fachübergreifende Unterrichtsforscher*innen zentrale Aspekte der Unterrichtsqualität definitorisch voneinander ab und bestimmen zugehörige Subdimensionen. In einem zweiten Schritt werden die Subdimensionen fachspezifisch in beobachtbare Indikatoren ausdifferenziert. Dabei werden Ähnlichkeiten und Unterschiede der fachspezifischen Konzeptionen über die Fächer verglichen, um ein einheitliches Verständnis von Unterrichtsqualität zu generieren.

Ergebnisse

In den ersten Austauschtreffen wurde deutlich, dass gewisse Dimensionen des Unterrichts (z. B. Klassenführung und sozio-emotionale Unterstützung) in allen Fächern wichtig sind, jedoch unterschiedliche Relevanz besitzen und auf Ebene der Subdimensionen unterschiedlich ausdifferenziert werden (z. B. Sicherheit im Fach Sport). Insbesondere der kognitiven Aktivierung wird in jedem Fach eine hohe Bedeutung zugemessen, wird jedoch in den Fachdidaktiken unterschiedlich definiert und erweitert. So wird beispielsweise im Fach Sport auch eine kognitiv-motorische Aktivierung (Herrmann & Gerlach, 2020) und im Literaturunterricht eine kognitiv-emotionale Aktivierung (Hesse & Winkler, 2022) vorgeschlagen. Entsprechend wurde diskutiert, ob und wie die Dimension Aktivierung lernziel- und gegenstandsorientiert ausdifferenziert werden kann. Es wird geprüft, ob der kognitiven Aktivierung weitere Typen der gegenstandsbezogenen Aktivierung zur Seite gestellt werde soll. Ziel dabei ist die Differenzierung des Konstrukts in Richtung einer lernzielbezogenen Fokussierung, wobei unterschiedliche Lerngegenstände auch unterschiedliche Arten der Aktivierung implizieren. Dabei soll die kognitive Aktivierung nicht beliebig breit definiert werden, sondern von weiteren lernzielorientierten Aktivierungsformen differenziert werden. Es wird angezielt, dass das Rahmenmodell über fachspezifische Ausdifferenzierungen verfügt und jenen (Sub-) Dimensionen der Unterrichtsqualität Raum gegeben wird, die sich auf den Aufbau, die Integration, Umstrukturierung, Erweiterung und Differenzierung fach- und konzeptspezifischer Lernprozesse richten (Reusser & Pauli, 2021).

Diskussion

Der zentrale Innovationsgehalt des Projekts liegt in der Verknüpfung von fachspezifischen und fachübergreifenden Konzeptionen von Unterrichtsqualität. Eine zentrale Entwicklungsperspektive auf konzeptioneller und auf empirischer Ebene besteht in der Entwicklung eines Modells, welches allgemeine Aspekte von Unterrichtsqualität darstellt und sich für verschiedene Fächer ausdifferenzieren lässt. Eine zweite Perspektive liegt auf der Ebene der Implementation. So wird sich der Erfolg dieses Vorhabens an der breiten Implementation in die Unterrichtspraxis messen lassen

müssen. Daraus ergibt sich methodisch der Anspruch, die konzeptionellen und empirischen Arbeiten in einem iterativen Prozess gemeinsam mit der Unterrichtspraxis zu entwickeln und zu erproben. Neben einem Rahmenkonzept sollen auch Evaluationsinstrumente und Implementationsmaterialien entwickelt werden, welche über die verschiedenen Fächer und Disziplinen als nützlich anerkannt werden.

Wie können innovative und belastbare Forschungsdesigns der Unterrichtsqualitätsforschung aussehen?

Richard Göllner[1], Aiso Heinze[2], Miriam Hess[3], Carmen Köhler[4], Wolfgang Wagner[1]
[1]Universität Tübingen, [2]IPN Kiel, [3]Universität Bamberg, [4]DIPF

Im Zuge der Forderung nach evidenzorientierter Praxis im Bildungsbereich (z.B. Bauer, Prenzel & Renkl, 2015) ist auch die Unterrichtsforschung gefordert. Insbesondere steht die empirische Forschung vor der zentralen Herausforderung, belastbare und vor allem über verschiedene Kontextbedingungen hinweg generalisierbare Nachweise zur Wirksamkeit eines qualitätsvollen Unterrichts zu liefern und daraus praxistaugliche Handlungsempfehlungen zur Erreichung eines qualitätsvollen Unterrichts bereitzustellen (z.B. Pant, 2014; Wilkes & Stark, 2022). Vor diesem Hintergrund wurde im Rahmen der Forschungswerkstatt auch die Weiterentwicklung von Forschungsdesigns in der empirischen Unterrichtsforschung diskutiert. Im Beitrag werden die Ergebnisse dieser Diskussion anhand von drei konkreten Ansätzen vorgestellt. Hierzu werden zunächst die gängigen Forschungsdesigns der empirischen Unterrichtsforschung dargestellt und damit verbundene Einschränkungen aufgezeigt. Inspiriert durch eine interdisziplinäre Forschungsperspektive werden anschließend die drei folgenden Erweiterungen anhand von Forschungsbeispiele präsentiert:

a) Erweiterung längsschnittlicher Untersuchungsdesigns: Inspiriert durch die Forschungspraxis im Bereich der Entwicklungs- und Persönlichkeitsforschung wird die Bedeutung von Transitionsphasen zur Identifikation von lehrkraftspezifischen Lerneffekten auf Seiten der Schülerinnen und Schüler herausgestellt. Gleichzeitig werden neuere analytische Strategien zur Auswertung entsprechender Längsschnittdaten präsentiert.

b) Einsatz von Unterrichtsskripts in der experimentellen Forschung: Im Zentrum steht die experimentelle Variation von Unterricht zur Ermittlung etwaiger Lerneffekte. Insbesondere die Variation von längeren Unterrichtssequenzen anhand von Skripts bietet eine Reihe von forschungsmethodischen und praxisrelevanten Potentialen.

c) Digitalisierte Unterrichtsforschung zur Prozessbeschreibung: Schließlich wird anhand modernerer Forschungstechnologien gezeigt, auf welche Weise Unterrichtseffekte im Sinne einer Prozessbeschreibung untersucht werden können. Als Beispiele dienen die Blickbewegungsmessungen im Klassenzimmer, die Verwendung einer virtuellen Klassenzimmersimulation sowie die prozessdiagnostische Analyse von Lehrersprache.

Session C

Einzelbeiträge

C-01: Digitales Lernen I

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S06 S00 B29
Chair der Sitzung: Nina Mulaimovic

Allein in einem Feld aus Kacheln: Herausforderungen von Online-Fortbildungen für Fortbildner:innen in der Lehrkräftebildung

Nina Mulaimovic[1], Christin Lucksnat[1], Rebecca Lazarides[1], André Meyer[1], Marc Kleinknecht[2], Dirk Richter[1]
[1]Universität Potsdam, Deutschland; [2]Leuphana Universität Lüneburg, Deutschland

Theoretischer Hintergrund

Die globale COVID-19-Pandemie erforderte eine kurzfristige Umstellung von analogen auf digitale Lehr-Lern-Umgebungen. Dass diese Umstellung sowohl Lernende als auch Lehrende an Schulen und Hochschulen vor zahlreiche Herausforderungen stellte und nach wie vor stellt, konnte bereits in verschiedenen Untersuchungen nachgewiesen werden (Garrote et al. 2021; Huber et al. 2020; Tengler et al. 2020; Traus et al. 2020). Beispielsweise wird von Schüler:innen und Studierenden häufig über unzureichende Austauschmöglichkeiten mit anderen Lernenden und Lehrenden berichtet (Garrote et al. 2021; Traus et al. 2020) und auch Letztere empfinden es als besonders herausfordernd, den persönlichen Kontakt zu ihren Lerngruppen aufrechtzuerhalten (Huber et al. 2020).

Lehrende in der Lehrkräftefortbildung wurden im Zuge der COVID-19-Pandemie ebenso wie Lehrende an Schulen und Hochschulen sehr kurzfristig mit der Aufgabe konfrontiert, ihre Präsenzveranstaltungen in ein digitales Format zu überführen (MBJS 2022; SiN 2020). Es liegt also nahe, dass diese ähnlichen Problemen ausgesetzt waren. Bislang liegen zu den Herausforderungen bei der Konzeption und Umsetzung digitaler Fortbildungen aber nur wenige Forschungsbefunde vor, die zudem meist nur die Wahrnehmung der Fortbildungsteilnehmer:innen widerspiegeln (Howard 2021; Wynants und Dennis 2018).

Fragestellung

Um Herausforderungen bei der Konzeption und Umsetzung digitaler Fortbildungen für Fortbildner:innen in der Lehrkräftebildung identifizieren zu können, verfolgt der vorliegende Beitrag primär die Forschungsfrage:

1. *Welche Herausforderungen bringen Online-Fortbildungen für Fortbildner:innen in der Lehrkräftefortbildung mit sich?*

Darüber hinaus wird untersucht, ob sich die wahrgenommenen Herausforderungen in Abhängigkeit bestimmter berufsbezogener Merkmale wie der Erfahrung der Fortbildner:innen unterscheiden. Dabei wird sowohl die allgemeine Berufserfahrung als Fortbildner:in als auch die spezifische Erfahrung mit der Durchführung von Online-Fortbildungen betrachtet, woraus die folgenden zwei Forschungsfragen resultieren:

1. *Unterscheiden sich die wahrgenommenen Herausforderungen in Abhängigkeit der Anzahl an Jahren, die die Fortbildner:innen bereits in diesem Beruf tätig sind?*
2. *Unterscheiden sich die wahrgenommenen Herausforderungen in Abhängigkeit der Anzahl an Online-Fortbildungen, die die Fortbildner:innen bereits durchgeführt haben?*

Methode

Die vorliegenden Daten basieren auf einer im Zeitraum von April bis Oktober 2021 durchgeführten, schriftlichen Online-Befragung mit 142 Fortbildner:innen in der Lehrkräftebildung. Im Zuge dessen wurden zum einen die von ihnen wahrgenommenen Herausforderungen digitaler Fortbildungsveranstaltungen und zum anderen ihre allgemeine Berufserfahrung als Fortbildner:in sowie ihre spezifische Erfahrung mit der Durchführung von Online-Fortbildungen erfasst.

Die Auswertung der Daten erfolgte mithilfe der qualitativen Inhaltsanalyse nach Mayring und Frenzl (2019).

Ergebnisse und ihre Bedeutung

Es konnten die drei Herausforderungskategorien *Beziehung zu und zwischen den Teilnehmenden (B)*, *Fortbildungsmodalität (M)*, und *Fortbildungsgestaltung (G)* identifiziert werden.

Am häufigsten wurden Schwierigkeiten berichtet, die sich der *Beziehung zu und zwischen den Teilnehmenden (B)* zuordnen lassen (69 %), wohingegen Herausforderungen, die die *Fortbildungsmodalität (M)* betreffen (50,4 %), am zweithäufigsten, und jene, die sich auf die *Fortbildungsgestaltung (G)* zurückführen lassen (44,2 %), am dritthäufigsten von den befragten Personen genannt wurden. Außerdem zeigt sich, dass sich Fortbildner:innen mit geringer Berufserfahrung signifikant von Fortbildner:innen mit hoher Berufserfahrung unterscheiden. So fühlen sich Erstere in Hinblick auf den Aufbau und die Aufrechterhaltung einer positiven Beziehung zu und zwischen den Teilnehmenden stärker herausgefordert als Letztere. Hingegen konnten für Fortbildner:innen mit unterschiedlicher Erfahrung in Online-Fortbildungen keine signifikanten Unterschiede ausfindig gemacht werden.

Aus den Ergebnissen der vorliegenden Studie lassen sich verschiedene Implikationen für Praxis und Forschung ableiten. Etwa empfiehlt es sich, Fortbildner:innen in der Lehrkräftebildung - vor dem Hintergrund, dass ihnen insbesondere der Aufbau und die Aufrechterhaltung einer positiven Beziehung schwerfällt - für beziehungsförderliche Maßnahmen zu sensibilisieren. Des Weiteren könnte es in Hinblick auf den Aufbau und die Aufrechterhaltung einer positiven Beziehung angemessen sein, vor allem Fortbildner:innen mit geringer Berufserfahrung zu unterstützen, da diese sich dahingehend besonders herausgefordert fühlen. Demzufolge erscheint es mit Blick auf zukünftige Forschung als besonders relevant, zu untersuchen, welche Unterstützungsmaßnahmen ergriffen werden können, um den von den Fortbildner:innen genannten Herausforderungen entgegentreten zu können.

Profile von Dozierenden während der Online-Lehre – eine Untersuchung auf Basis von digitalen Lernaktivitäten

Anne Lohr, Michael Sailer, Frank Fischer
Ludwig-Maximilians-Universität München (LMU), Deutschland

Wenn Dozierende an Hochschulen digitale Medien effektiv in Lehrveranstaltungen einsetzen, kann das die Kompetenzen der Studierenden fördern (Sailer et al., 2021). Da die reine Häufigkeit des Medieneinsatzes durch die Dozierenden nicht per se mit besseren Lernergebnissen der Studierenden in Zusammenhang steht (Tamim et al., 2011), wird in dieser Studie ein Fokus auf die kognitive Aktivierung der Studierenden während digitaler Lernaktivitäten vorgeschlagen. Auf Basis des ICAP-

Modells (Chi & Wylie, 2014) können vier verschiedene Arten von digitalen Lernaktivitäten unterschieden werden: Passiv (P), aktiv (A), konstruktiv (K) und interaktiv (I). In der Präsenzlehre konnten drei Profile von Dozierenden hinsichtlich ihrer Initiierung dieser vier digitalen Lernaktivitäten gefunden werden: (a) Die „Powerpointers“, die hauptsächlich passive Lernaktivitäten, (b) die „Clickerers“, die hauptsächlich passive und aktive Lernaktivitäten und, (c) die „Digital Pros“, die alle vier Lernaktivitäten etwa zu gleichen Teilen initiieren (Lohr et al., 2021).

Ziel dieser empirischen Studie ist es zu untersuchen, inwiefern sich diese drei gefundenen Profile aus der Präsenzlehre im Kontext der Online-Lehre während der Corona-Pandemie replizieren lassen. Darüber hinaus soll untersucht werden, inwiefern sich die Dozierenden in den verschiedenen Profilen bezüglich ihrer medienbezogenen Kompetenzen (Technologische Kenntnisse, TK; Technologisch-pädagogische Kenntnisse, TPK) und ihren Einstellungen (Allgemeine Einstellungen zu digitalen Medien; Einstellungen zu digitalen Medien in der Lehre) unterscheiden.

Die Daten von N = 1173 Dozierenden staatlicher bayerischer Hochschulen wurden mit einer Online-Befragung Ende 2021 erhoben. Die Dozierenden sollten sich in ihren Antworten allerdings auf das vorausgegangene, von der Corona-Pandemie geprägte Sommersemester 2021 beziehen. 43.10 % der Befragten waren weiblich und waren im Durchschnitt 44 Jahre alt. Die vier Arten digitaler Lernaktivitäten wurden auf einer 5-stufigen Likert-Skala gemessen, die medienbezogenen Kompetenzen mittels objektiver Wissenstests (Kastorff et al., 2022; Lachner et al., 2019) und die medienbezogenen Einstellungen mittels der Skala von Fleck et al. (2021).

Anhand einer explorativen k-means Clusteranalyse zeigen sich vier Profile der Dozierenden: Während sich die Profile der Powerpointers (N = 204) und der Clickerers (N = 341) identifizieren ließen, zeigten sich zwei weitere Profile: die Orchestrators (N = 307), welche alle Arten von digitalen Lernaktivitäten vergleichbar häufig einsetzen, und die „Interactivators“ (N = 231), welche vor allem digitale Lernaktivitäten am oberen Ende des ICAP-Kontinuums und dort insbesondere interaktive Lernaktivitäten einsetzen.

Ergebnisse von ANOVAs zeigen auf, dass bezogen auf technologische Kenntnisse keine signifikanten Unterschiede zwischen den vier Profilen gefunden werden können, $F(3,1079) = 1.20$, $p = .310$, $\eta_p = 0.003$. Bezogen auf TPK unterscheiden sich die Profile jedoch signifikant voneinander, $F(3,1079) = 19.30$, $p < .001$, $\eta_p = 0.051$. Powerpointers und Clickerers weisen jeweils signifikant niedrigeres TPK auf als die Orchestrators und die Interactivators. Bezogen auf allgemeine Einstellungen zu digitalen Medien, $F(3,1078) = 20.80$, $p < .001$, $\eta_p = 0.055$, und Einstellungen zu digitalen Medien in der Lehre, $F(3,1079) = 19.00$, $p < .001$, $\eta_p = 0.050$, zeigten sich ebenfalls signifikante Unterschiede. Die Befundmuster der beiden Einstellungsskalen sind hierbei sehr ähnlich: Powerpointers haben kritischere Einstellungen als alle anderen Profile, Orchestrators haben positivere Einstellungen als alle anderen Profile. Clickerers und Interactivators unterscheiden sich nicht signifikant voneinander und befinden sich im Mittelfeld zwischen den beiden anderen Profilen, von denen sie wiederum jeweils signifikant abweichen.

Die Profile aus der Präsenzlehre konnten somit teilweise repliziert werden – es kommt allerdings in der Online-Lehre zu einer Ausdifferenzierung der Digital Pros in Form der Orchestrators und

Interactivators. Beide Profile initiieren häufiger die konstruktiven und interaktiven Lernaktivitäten, was somit mit höherer Wahrscheinlichkeit zu höherem Kompetenzerwerb der Studierenden führt. Für diese Art der Lehre scheint das TPK der Dozierenden möglicherweise eine wichtige Rolle spielen, das diese beiden Profile hier ausgeprägtere Kenntnisse vorweisen. Zukünftig soll untersucht werden, wie es zu den beiden effektiveren Profilen kommt, beispielweise durch das Vorhandensein einer Digitalisierungsstrategie oder Weiterbildungen.

Zusammenhang von Ungewissheitstoleranz von Lehrpersonen und Nutzung digitaler Medien im Unterricht

Julia Gerick[1], **Dagmar Killus**[2]
[1]TU Braunschweig, Germany; [2]Universität Hamburg, Germany

Theoretischer Hintergrund

Die fortschreitende Digitalisierung aller Lebens- und Arbeitsbereiche stellt Schulen vor große Entwicklungsaufgaben. Das hängt insbesondere damit zusammen, dass mit den rasanten technologischen Entwicklungen immer auch neue Möglichkeiten für das Lehren und Lernen einhergehen.

Dass die Ungewissheitstoleranz von Lehrpersonen deren Umgang mit beruflichen Anforderungen positiv beeinflusst, gilt als empirisch gesichert (z.B. König & Dalberg 2007). Es wurden jedoch bislang nur wenige systematische Anstrengungen unternommen, um zu verstehen, wie die Ungewissheitstoleranz mit (weiteren) internen Ressourcen und externen Ressourcen zusammenhängt und welche Auswirkungen sich daraus für den Einsatz digitaler Medien im Unterricht ergeben. Mit *lehrerbezogenen Einstellungen* (Selbstwirksamkeitsüberzeugungen und Sichtweise auf den Nutzen digitaler Medien) und mit *Lehrerkooperation* auf der Ebene der Schule werden diejenigen internen und externen Ressourcen berücksichtigt, die nachweislich auch zentrale Bedingungsfaktoren für die Nutzung digitaler Medien im Unterricht darstellen (Drossel et al. 2019; Tondeur et al. 2019; Woltran et al. 2022).

Fragestellung

Gestützt auf Daten der international vergleichenden Schulleistungsstudie ICILS 2018 (International Computer and Information Literacy Study; Eickelmann et al., 2019) geht der Beitrag drei Fragestellungen nach:

(1) Welche Zusammenhang besteht zwischen der Ungewissheitstoleranz von Lehrpersonen und der Form der Nutzung digitaler Medien im Unterricht?

(2) Welche internen Ressourcen von Lehrpersonen (Selbsteinschätzung digitalisierungsbezogener Kompetenzen, positive Sichtweise auf digitale Medien) und welche externen Ressourcen ihres Arbeitsplatzes (Lehrerkooperation) weisen einen Zusammenhang mit der Ungewissheitstoleranz auf?

(3) Wie wirken sich Ungewissheitstoleranz, interne und externe Ressourcen gemeinsam auf die Form der Nutzung digitaler Medien im Unterricht aus?

Methode

Den Fragestellungen wird auf Grundlage der Lehrpersonendaten aus der Studie ICILS 2018 mittels Regressionsanalyse (Forschungsfragen 1 und 2) sowie Strukturgleichungsmodellierung (Forschungsfrage 3) mit Mplus nachgegangen. Entsprechend der Zielpopulation von ICILS 2018 umfasst die Datengrundlage 2328 Lehrpersonen in Deutschland, die in der achten Jahrgangsstufe unterrichten, an insgesamt 182 Schulen. Der Komplexität der Datenstruktur wird mit type=complex sowie dem robusten MLR-Schätzer Rechnung getragen.

Ergebnisse und ihre Bedeutung

Die Befunde bestätigen die Annahme, dass die Ungewissheitstoleranz von Lehrpersonen auch mit dem Einsatz digitaler Medien im Unterricht eng zusammenhängt: Je ungewissheitstoleranter Lehrpersonen sind, desto stärker nutzen sie in ihrem Unterricht digitale Medien. Die Ungewissheitstoleranz von Lehrpersonen hängt (erwartungsgemäß) mit internen und externen Ressourcen zusammen. Als besonders erklärungsstark haben sich hier die Sichtweise von Lehrpersonen auf den Nutzen digitaler Medien und die Lehrerkooperation erwiesen. Auch bei statistischer Kontrolle konkurrierender interner Ressourcen und externer Ressourcen leistet Ungewissheitstoleranz einen eigenständigen Beitrag zur Erklärung des Umgangs mit beruflichen Anforderungen – hier der Einsatz digitaler Medien (König & Dalbert 2007; König & Dalbert 2004). Welche Konsequenzen sich aus den Ergebnissen für die Professionalisierung von Lehrpersonen in und für Schulentwicklungsprozesse ableiten lassen, soll diskutiert werden.

C-02: Einfluss der Eltern

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S06 S00 B32
Chair der Sitzung: Kerstin Schörner

Students' perceptions or parents' factual aspirations - the potential problem of reverse causality in operationalizations of the Wisconsin Model

Kerstin Schörner, Felix Bittmann
Leibniz Institute for Educational Trajectories, Germany

Aspirations have been shown to be crucial for the explanation of educational inequalities (Ditton, Krüsken, & Schauenberg, 2005). Therefore, science and politics need to understand their development. One prominent sociological theory aiming to explain the development of aspirations as dependent on social background is the Wisconsin Social Psychological Model of Status Attainment (Sewell, Haller, & Portes, 1969). A central assumption of the model suggests that students' perception of their parents' aspirations shapes students' aspirations, which then determines students' educational and occupational attainment (Sewell et al., 1969; Sewell & Hauser, 1980). The authors state clearly that it is students' perception that is crucial for the development of their aspirations and not the objectively measured expectations parents state (Sewell & Hauser, 1980; Sewell & Shah, 1968). In contrast to this, literature investigating the Wisconsin Model often operationalizes parents' aspirations as measured by parents instead of taking students' perceptions. The main argument for doing so is the possible problem of reverse causality. Instead of students' perceptions of their parents' aspirations influencing students' aspirations, it might be the other way around and students' own aspirations are projected on their perceptions of what their parents wish for (Hoenig, 2019; Zimmermann, 2018). In our paper, we ask if the substitution of students' perception of their parents' aspirations with the parents' aspirations, directly asked from them, is reasonable and valid for the investigation of the Wisconsin Models' mechanism.

Two research questions are addressed. First, we ask if the potential problem of reverse causality, namely the projection of students' own aspirations on their perception of their parents' aspirations, can be proven empirically. Second, we investigate whether the aspirations measured by parents can explain as much as the student's perception of parents' aspirations can do. NEPS data from starting cohort 3, wave 2, 4 and 5, are used for analysis (Blossfeld & Roßbach, 2019). Aspirations are measured using a binary construct, indicating whether children have *idealistic* aspirations for the *Abitur* (1) or not (0) and whether they believe that their parents have these aspirations for them or not. To avoid spurious correlations a large set of relevant control variables, especially for social origin, is included in all models. Concerning research question 1, cross lagged panel models (CLPM) with three measurement points show statistically significant crossed effects: from students' aspiration to students' perception of their parents' aspirations and vice versa. Thereby, the problem of reverse causality is empirically detected. When effect sizes are compared it turns out that the effect of own aspirations on perceived aspirations is stronger, sometimes almost twice the effect size as in the other direction.

Regarding research question 2, coefficients of linear fixed effect regressions and within-person R^2 show students' perceptions of parents' aspirations to be more relevant for explaining their aspirations than factual parental aspirations. As a conclusion one can state that, in order to make causal inferences, one should avoid taking students' perceptions of parents' aspirations since we empirically

proofed the assumed problem of reverse causality. Students do indeed project their own aspirations on the perception of their parents' aspirations. However, using parental aspirations as measured in parents' interviews, researchers should consider their analysis to underestimate the influence parents' aspirations have through the perception of their children.

Mikroanalytische Messung elterlicher Kontingenz: Kodierschema für Eltern-Kind-Interaktionen während des gemeinsamen Problemlösens im Kindergartenalter

Leonie Köllbichler, Maren Stern, Silke Hertel
Universität Heidelberg, Germany

Theoretischer Hintergrund

Elterliche Ko-Regulation stellt einen essentiellen und gestaltenden Faktor in der Selbstregulationsentwicklung im Kindesalter dar und gilt als bedeutsamer Resilienzfaktor bei familiären Risiken (Fay-Stammbach et al., 2014; Valcan et al., 2017). Ko-regulative Verhaltensweisen wie einfühlsame Reaktionen auf die Bedürfnisse des Kindes oder kognitive Unterstützung in Problemlösesituationen helfen dem Kind sich zu regulieren und werden zunehmend vom Kind internalisiert (Bernier et al., 2010). Eine Vielzahl von Studien identifizieren Kontingenz als essentielle Komponente förderlicher Ko-Regulation für die Entwicklung von Selbstregulation (Kallia et al., 2022; Mermelshtine, 2017).

Kontingente Unterstützung richtet sich nach dem Fähigkeitslevel des Kindes: Nach einem Misserfolg während einer gemeinsamen Aufgabe wird das Ausmaß an Unterstützung erhöht und nach einem Erfolg verringert (Wood et al. 1978). Kontingente Unterstützung ist hoch dynamisch und wurde bisher selten auf mikroanalytischer Ebene erfasst. Eine mikroanalytische Erfassung ermöglicht es jedoch diese Dynamik abzubilden (Pino-Pasternak et al., 2010). Im Vorschulalter spielt die nonverbale Ebene eine wesentliche Komponente in der Interaktion (Siegler et al., 2021). Bisherige Kodierschemata für Eltern-Kind-Interaktionen mit älteren Kindern fokussierten stärker auf das verbale Verhalten (siehe Hermkes, et al. 2018; Pino-Pasternak et al., 2010). Diese Studie hat das Ziel ein mikroanalytisches Kodierschema für die Erfassung elterlicher Kontingenz in Interaktion mit ihrem Kind im Kindergartenalter zu entwickeln.

Fragestellung und Ziele

Wie kann elterliche Kontingenz mikroanalytisch während kollaborativer Eltern-Kind-Interaktionen mit Kindern im Kindergartenalter gemessen werden? Zur Beantwortung dieser Frage werden zwei Ziele verfolgt:

(1) ein neues Kodierschema zu erstellen, das die Erfassung des nonverbalen Verhaltens des Kindes fokussiert, und

(2) das resultierende Kodierschema auf seine Reliabilität und Validität zu prüfen.

Methode

In Anlehnung an Pino-Pasternak et al. (2010) wurde ein mikroanalytisches Kodierschema entwickelt, welches drei Ebenen unterscheidet: 1) lösungsorientiertes Verhalten des Kindes (0=nicht lösungsorientiert bis 4=löst die Aufgabe), 2) Unterstützungsgrad des Elternteils (0=keine aktive Unterstützung bis 4=löst die Aufgabe) und 3) relative Kontingenz, basierend auf der Passung von

Eltern- und Kindverhalten (kontingent, nicht kontingent: zu viel, zu wenig oder keine Unterstützung). Zur mikroanalytischen Erfassung wird das Verhalten jeweils in 5-Sekunden-Intervallen kodiert.

Das Kodierschema wurde an einem Datensatz von *N*=30 Mutter-Kind-Dyaden (*M* = 4.39 Jahre, *SD* = 0.31) erprobt. Im Rahmen der Studie bearbeiteten Mutter-Kind-Dyaden eine räumlich-visuelle Problemlöseaufgabe (KABC-II, Melchers & Melchers, 2015). Diese Interaktionssequenz wurde videographiert und mit dem neu entwickelten Kodierschema analysiert. Zudem wurde die globale Kontingenz (Erdmann et al., 2019) und Problemlöseleistung während der gemeinsamen Bearbeitung erfasst.

Ergebnisse

Zur Beurteilung der Reliabilität des Kodierschemas wurden Interrater-Reliabilitäten basierend auf 10% der Videos berechnet. Es zeigten sich ausgezeichnete Übereinstimmungen für das lösungsorientierte Verhalten des Kindes (ICC = 0.98) und gute Übereinstimmungen für den Unterstützungsgrad der Mutter (ICC = 0.77)). Zur Validierung des Schemas wurden Korrelationen zwischen der relativen Kontigenz mit 1) der globalen Kontingenz und 2) der Problemlöseleistung gerechnet. Die ermittelte relative Kontingenz korrelierte nicht signifikant mit der globalen Kontingenz (=.186, *p*=.343) und der Problemlöseleistung (*r*=.213, *p*=.277).

Das Kodierschema ermöglicht unterschiedliche mikroanalytische Untersuchungen von Eltern-Kind-Interaktionen: Einschätzungen der elterlichen Kontingenz für ganze Interaktionssequenzen, Formen nicht-kontingenten Verhaltens zu untersuchen sowie einzelne Ausschnitte aus einer Interaktionssequenz zu analysieren. Aus der Kodierung resultiert zusätzlich eine graphische Abbildung der Verhaltens- und Kontingenzebenen, die für qualitative Untersuchungen oder in der praktischen Arbeit mit Eltern genutzt werden kann. Diese verschiedenen Herangehensweisen wurden an ausgewählten Fallbeispielen exploriert.

Bei der Interpretation der Ergebnisse ist die kleine, selektive Stichprobengröße zu berücksichtigen, jedoch wird das Kodierschema noch an weiteren Datensätzen erprobt. Insgesamt zeigt das neu entwickelte Kodierschema vielversprechende Nutzbarkeit in vielen Bereichen von Wissenschaft und Praxis, die Validität muss noch in einer größeren, heterogeneren Stichprobe überprüft werden.

Wortanzahl: 570

Zur Erklärung von Unterschieden im entscheidungsrelevanten Bildungswissen zwischen Eltern mit und ohne Migrationshintergrund in Deutschland

Thomas Zimmermann
Johann Wolfgang Goethe-Universität, Germany

Theoretischer Hintergrund und Fragestellung

Informationen wirken sich auf die Bildungsbeteiligung von Familien mit einem Zuwanderungshintergrund und damit auf ihre strukturelle Integration in die Aufnahmegesellschaft aus (Diehl, Hunkler und Kristen, 2016). Insbesondere am Übergang in die Sekundarstufe I sind Informationen von Bedeutung, denn an diesem institutionell definierten Übergang treffen Familien eine Entscheidung über den weiteren Bildungserfolg ihrer Kinder (Maaz *et al.,* 2008). Trotz der Relevanz von Informationen haben nur wenige Studien untersucht, ob Unterschiede in der

Verfügbarkeit von Informationen zwischen Personen mit und ohne Zuwanderungshintergrund bestehen und wie sich diese erklären lassen (vgl. Kretschmer, 2019). Eine Erklärung von Informationsunterschieden bieten verschiedene Theorien. Zentral ist die Bedeutung familiärer Ressourcen (Bourdieu, 2012). Bourdieu (2012) unterscheidet drei für den Bildungserfolg entscheidende Ressourcen: ökonomisches, soziales und kulturelles Kapital, wobei soziales und kulturelles Kapital eine zentrale Rolle bei der Erklärung von Informationsunterschieden spielen. Das Ausmaß, in dem die Ausstattung von Eltern mit Zuwanderungshintergrund mit sozialem und kulturellem Kapital zu ihrem Informationsbestand beiträgt, hängt von der Instrumentalität ihres Kapitals (Friedberg, 2000) und ihren Akkulturationsstrategien ab (Berry, 1997). Unsere Fragestellung lautet:

Welche Unterschiede bestehen in der Bildungsinformiertheit zwischen Personen mit und ohne Zuwanderungshintergrund und wie lassen sich diese erklären?

Daten und Methoden

Datengrundlage ist die Startkohorte 2 des Nationalen Bildungspanel (NEPS-Netzwerk 2021). Für den Übergang von der Primar- zur Sekundarstufe I stehen Daten von insgesamt 3.207 Eltern sowie Schülerinnen und Schülern zur Verfügung. Das NEPS ist für die Beantwortung unserer Forschungsfrage gut geeignet, denn es bestehen zum einen grundlegende Daten über die Bildungsinformiertheit der Eltern wie Ihr Wissen zu Übergangsregelungen. Zum anderen bietet das NEPS reichhaltige Informationen über relevante Erklärungsfaktoren wie die sozialen (Positionsgenerator) und kulturellen Ressourcen (Bildungshintergrund und Sprachkenntnisse), die den Eltern zur Verfügung stehen. Darüber hinaus sind differenzierte Analysen nach dem Zuwanderungshintergrund möglich (Olczyk, Will und Kristen, 2014). Mithilfe von linearen Regressionsanalysen untersuchen wir Unterschiede in der Informationsverfügbarkeit zwischen Eltern mit und ohne Zuwanderungshintergrund. Zur Erklärung der Unterschiede verwenden wir entsprechende Mediationsanalysen.

Ergebnisse und ihre Bedeutung

Unsere Analysen bestätigen, dass Eltern mit einem Zuwanderungshintergrund über eine geringere Bildungsinformiertheit verfügen als Eltern ohne einen Zuwanderungshintergrund. Darüber hinaus zeigen sich Unterschiede in der Informiertheit nach Zuwanderungsgruppen. Im Vergleich zu Eltern ohne Zuwanderungshintergrund ist insbesondere bei Eltern aus der Türkei und Nordafrika eine geringere Bildungsinformiertheit feststellbar. Unsere Mediationsanalysen zeigen, dass Unterschiede in der Bildungsinformiertheit zwischen den verschiedenen Zuwanderungsgruppen weitestgehend durch eine unterschiedliche Ausstattung mit sozialem und kulturellem Kapital erklärbar sind: Wissensunterschiede nehmen mit steigender Bildung und sozialer Herkunft, mit einer zunehmenden Nutzung der Verkehrssprache, mit eigenen Bildungserfahrungen in Deutschland und mit einem ansteigenden Anteil hochgebildeter ohne Zuwanderungshintergrund in dem sozialen Netzwerk ab.

Obwohl der Informiertheit von Eltern mit einem Zuwanderungshintergrund eine bedeutsame Rolle für die Bildungsbeteiligung ihrer Kinder und damit auch für die strukturelle Integration im Aufnahmeland beigemessen wird, wissen wir bisher wenig darüber, wie groß diese Informationsunterschiede sind

und wie sie sich erklären lassen. Die vorliegende Arbeit füllt diese Forschungslücke, indem sie das Ausmaß der Unterschiede im Bildungswissen zwischen Personen mit und ohne Zuwanderungshintergrund aufzeigt und erklärt. Wenn das Ziel ist, dass Eltern mit einem Zuwanderungshintergrund strategisch „richtige“ Entscheidungen treffen (vgl. Diehl, Hunkler und Kristen, 2016) dann müssen Sie zunächst über eine geeignete Informationsgrundlage verfügen, die Ihnen eine solche Entscheidungsfindung ermöglicht.

Zusammenhänge zwischen elterlicher Hausaufgabenhilfe, Hausaufgabenverhalten und schulischer Leistung – Spielt das Alter der Schüler:innen eine Rolle?

Lisa Benckwitz[1], Katharina Kohl[1], Natalia Suárez[2], José Carlos Núñez[2], Karin Guill[3]
[1]Leibniz Institute for Science and Mathematics Education Kiel, Deutschland; [2]University of Oviedo; [3]Ständige Wissenschaftliche Kommission der Kultusministerkonferenz

Theoretischer Hintergrund

Die bisherige Forschung zu Effekten elterlicher Hausaufgabenhilfe hat gezeigt, dass die Qualität dieser Unterstützungsform einen wichtigeren Prädiktor für das Hausaufgabenverhalten und die Leistung von Schüler:innen darstellt als die Quantität (Dumont et al., 2014; Moroni et al., 2015). Um die Qualität elterlicher Hausaufgabenhilfe zu messen, haben Dumont et al. (2014) auf Grundlage der Selbstbestimmungstheorie (Ryan & Deci, 2000) drei Qualitätsdimensionen vorgeschlagen: Ansprechbarkeit (z.B. helfen, wenn das Kind darum bittet), Struktur (z.B. Herstellen einer ruhigen Hausaufgabenumgebung) und Kontrolle (z.B. helfen, obwohl das Kind keine Hilfe benötigt). Für Ansprechbarkeit und Struktur sind theoretisch positive Zusammenhänge mit Schüler:innen-Outcomes zu erwarten, da sie zur Erfüllung der psychologischen Grundbedürfnisse nach Kompetenz, Autonomie und sozialer Verbundenheit beitragen und somit die intrinsische Motivation fördern können. Für Kontrolle hingegen werden negative Zusammenhänge erwartet, da kontrollierendes Elternverhalten die Erfüllung der Grundbedürfnisse behindern kann. Diese theoretisch abgeleiteten Zusammenhänge konnten bereits empirisch bestätigt werden (Dumont et al., 2014; Guill et al., 2020; Moroni et al., 2015). Es gibt jedoch erste Hinweise, dass die Effekte der Hausaufgabenhilfe möglicherweise altersspezifisch sind (Núñez et al., 2015; Pattall et al., 2008). Allerdings haben bisher nur wenige Studien die Zusammenhänge zwischen der Qualität elterlicher Hausaufgabenhilfe und den Outcomes von Schüler:innen in unterschiedlichen Altersgruppen miteinander verglichen (Gonida & Cortida, 2014; Núñez et al., 2015).

Daher adressiert die aktuelle Studie die Forschungsfrage, ob sich die Beziehungen der Qualitätsdimensionen der elterlichen Hausaufgabenhilfe mit dem Hausaufgabenverhalten und der Leistung von Schüler:innen zwischen verschiedenen Altersgruppen unterscheiden.

Methode

Zur Beantwortung der Fragestellung wurde eine Stichprobe von *N* = 1782 Schüler:innen (55% weiblich) der Klassen 5 bis 9 herangezogen. Die Schüler:innen beurteilten die Qualität ihrer elterlichen Hausaufgabenhilfe hinsichtlich Ansprechbarkeit, Struktur und Kontrolle (14 Items, Dumont et al., 2014) sowie ihr Hausaufgabenverhalten (vier Items für *Anstrengungsbereitschaft*, adaptiert nach Trautwein et al., 2006; drei Items für *Prokrastination*, Dumont et al., 2014) und gaben ihre letzten Zeugnisnoten in Mathematik, Deutsch und Englisch an, die wir umgepolt und als Indikator für

die Leistung als Durchschnittsnote zusammengefasst haben. In den Analysen haben wir für das Geschlecht sowie den Familien- und Bildungshintergrund kontrolliert.

Angelehnt an Núñez et al. (2015) wurde die Stichprobe in drei Substichproben aufgeteilt ($n_{Klassen5/6}$ = 721, $n_{Klassen7/8}$ = 694, $n_{Klasse9}$ = 312), um latente Mehrgruppenanalysen in M*plus* 8.5 durchzuführen. Fehlende Werte wurden mittels FIML behandelt.

Ergebnisse und ihre Bedeutung

Wir fanden altersspezifische Unterschiede in den Beziehungen zwischen den Qualitätsdimensionen und den Outcomes der Schüler:innen, die Ergebnisse waren jedoch komplex, da sie sich zwischen den Qualitätsdimensionen und Outcomes unterscheiden. Ansprechbarkeit und Struktur hingen fast durchgängig signifikant positiv mit der Anstrengungsbereitschaft zusammen (.13 ≤ β ≤ .46, $p \leq .05$). Lediglich in der mittleren Altersgruppe (Klasse 7 und 8) gab es keinen signifikanten Zusammenhang zwischen Ansprechbarkeit und Anstrengungsbereitschaft. Ansprechbarkeit und Struktur hingen jeweils in den zwei jüngeren Altersgruppen – nicht aber in der ältesten Gruppe (Klasse 9) – signifikant mit Prokrastination zusammen, wobei die negativen Koeffizienten eine Abnahme der Prokrastination abbilden (–.32 ≤ β ≤ –.11, $p \leq .05$). Entgegen der Erwartung hingen Ansprechbarkeit und Struktur in den zwei jüngeren Stichproben teilweise mit schlechteren Durchschnittsnoten zusammen (–.16 ≤ β ≤ –.06, $p \leq .05$).

Kontrolle zeigte ungünstige Zusammenhänge mit dem Hausaufgabenverhalten in den zwei jüngeren Gruppen (Anstrengungsbereitschaft: $\beta_{Klassen5/6}$ = –.28, $p \leq .001$; $\beta_{Klassen7/8}$ = –.23, $p \leq .001$; Prokrastination: $\beta_{Klassen5/6}$ = .47, $p \leq .001$; $\beta_{Klassen7/8}$ = .24, p = .003), nicht aber in der ältesten Gruppe. Lediglich in der ältesten Gruppe hing Kontrolle signifikant mit schlechteren Durchschnittsnoten zusammen (β = –.16, p = .006).

Zusammengefasst zeigten Ansprechbarkeit und Struktur tendenziell in allen drei Altersgruppen förderliche Zusammenhänge mit dem Hausaufgabenverhalten, während Kontrolle altersspezifisch mit unterschiedlichen Outcomes zusammenhing.

C-03: Kompetenzentwicklung und -messung

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S06 S00 B41
Chair der Sitzung: Fabio Fortunati

Instruktionssensitivität als Indiz für heterogene Kompetenzentwicklung in der ökonomischen Bildung: Schulformbezogene DIF-Analysen am Beispiel der Jahrgangsstufe 8 in Nordrhein-Westfalen

Fabio Fortunati, Esther Winther
Universität Duisburg-Essen, Germany

In der ökonomischen Bildung in Deutschland treffen Schüler*innen auf höchst heterogene Lerngelegenheiten, was das Stundenkontingent, Fächerstruktur sowie die Lehrkräfteausbildung anbelangt (Fortunati, Winther 2021; Macha 2019; Siegfried, Hangen 2020). Befunde aus diversen Studien zum Einfluss von formalen schulischen Lerngelegenheiten auf die Entwicklung ökonomischer Kompetenz können nicht nur deshalb eingeschränkt miteinander verglichen werden; dies ist auch den unterschiedlichen methodischen Ansätzen sowie den differenten theoretischen Konzeptionen geschuldet (Siegfried, Wuttke 2021). Gerade deshalb scheint die Rolle individueller Merkmale als Prädiktor ökonomischer Kompetenz in Studien deutlich häufiger untersucht zu werden als institutionelle und curriculare Rahmenbedingungen (Siegfried, Ackermann 2020; Kaiser et al. 2020; Seeber et al. 2018). Erste Befunde zu institutionellen und curricularen Kontextfaktoren als Prädiktoren für ökonomische Kompetenz beziehen sich auf die Sekundarstufe II und vergleichen die gymnasiale und beruflich-gymnasiale Oberstufe. Wider Erwarten erreichen Gymnasiasten an allgemeinbildenden Schulen einen höheren Grad an ökonomischer Kompetenz (Siegfried, Ackermann 2020). Dies wirft unmittelbar die Passung zwischen Curriculum und Assessment, aber auch zwischen erfahrener Instruktion und Assessment auf. So können Inhaltsbereiche eines Assessments sich instruktionssensitiv auf spezifische Gruppen auswirken, da diesen mehr Lernmöglichkeiten zur Verfügung standen (Naumann et al. 2019a). Generell kann unter dem Konzept der Instruktionssensitivität verstanden werden, ob Lernerfolge unterrichtsbezogen attribuiert werden können (Polikoff 2010).

Ziel dieses Beitrags ist es, das Konstrukt Instruktionssensitivität für den Bereich der ökonomischen Bildung zu konzeptionalisieren. Hier soll vor allem die Bedeutung von kontextbezogenen Merkmalen wie die Schulform in den Blick genommen werden. Die Schulform kann hier als Proxy für unterschiedliche unterrichtsbezogenen sowie anderen institutionellen Einflussfaktoren betrachtet werden. Gleichwohl ist anzumerken, dass für eine tiefer gehende Analyse von Unterrichtseffekten der Einbezug relevanter unterrichtsbezogener Merkmale unerlässlich ist (Naumann et al. 2019b). Auf Basis der Feldtestdaten des Assessments ECON2022 von Schüler*innen der Jahrgangsstufe 8 in Nordrhein-Westfalen soll geprüft werden, ob das entwickelte Instrument (1) instruktionssensitiv bezogen auf die Schulform ist und (2) welche inhalts- und kognitionsbezogenen Merkmale diese Items aufweisen.

In der Literatur lassen sich die diversen Ansätze zur Messung von Instruktionssensitivität unter zwei Überschriften subsumieren: Der Judgmental-Approach ist üblicherweise in den Prozess des Testdesigns integriert, indem ex ante Aussagen über die Instruktionssensitivität von Items getroffen werden und auf Grundlage dessen Items konstruiert werden (Popham 2007). Der zweite Ansatz stellt die empirische Ex post-Untersuchung von Testergebnissen auf Basis der IRT-basierten Differential-

Item-Functioning dar (Naumann et al. 2019a; Masters 1988). Kernelement DIF-basierter Studien ist es, dass bei Schüler*innen mit gleicher Personenfähigkeit aufgrund eines gruppenbezogenen Merkmals unterschiedliche Lösungswahrscheinlichkeiten für das Assessment als Ganzes und auf Itemebene auftreten. In diesem Beitrag wird der Judgmental Approach mit dem empirischen Ansatz verbunden, in dem ex ante Hypothesen über mögliche instruktionssensitive Inhalte in den verschiedenen Schulformen formuliert und im Anschluss empirisch geprüft werden.

Das Sample umfasst N=817 Schüler*innen der Jahrgangsstufe 8 aus Nordrhein-Westfalen, deren ökonomische Kompetenz mithilfe eines authentisch konstruierten und computergestützten Assessment erhoben wurde. Die Ergebnisse der DIF-Analyse lassen sowohl auf Test- als auch auf Itemebene signifikante Schulformunterschiede erkennen. Auf Grundlage der curricularen Annahmen zur Instruktionssensitivität lassen sich aus den Ergebnissen inhalts- wie kognitionsbezogene Unterschiede feststellen. Diese Ergebnisse zeigen, dass aus inhaltlicher Perspektive Realschüler*innen und mit Einschränkungen auch Gesamtschüler*innen deutlich häufiger mit betriebswirtschaftlichen Inhalten in Berührung kommen als Gymnasiast*innen. Im Vergleich dazu zeigen sich Items rund um die Inhalte der Verbraucherbildung aus Konsumentensicht sowie der Nachhaltigkeit für die Gymnasiast*innen itemsensitiv. Aus kognitionsbezogener Perspektive zeigt sich, dass die Items, die prozedurales Wissen erfordern, sowie das Argumentieren und Begründen systematisch instruktionssensitiv für Gymnasiast*innen sind. Für die Real- und Gesamtschüler*innen erweisen sich Item, die deklarative Wissensbestände adressieren, als instruktionssensitiv. Die Ergebnisse können zum einen dabei helfen, instruktionssensitive Items für Assessments in der ökonomischen Bildung zu konstruieren und zum anderen für die Fort- und Weiterentwicklung intendierter und implementierter Curricula im Sinne instruktionaler Kohärenz genutzt werden.

Kompetenzentwicklung in beruflicher Bildung: Eine deskriptive Analyse über Stagnation oder Wachstum in dualen und schulischen Ausbildungen

Jacqueline Kroh, Markus Nester, Ilka Wolter
Leibniz-Institut für Bildungsverläufe, Germany

Ein zentrales Ziel von Bildung ist es, Fähigkeiten und Kenntnisse zu vermitteln, die für eine erfolgreiche Integration in den Arbeitsmarkt erforderlich sind. Je nach Struktur des Bildungssystems ist diese Aufgabe auch ein Ziel der beruflichen Bildung, die die Lücke zwischen der allgemeinen Bildung und dem Arbeitsmarkt schließt. Aufgrund des raschen Wandels beruflicher Anforderungen und des Arbeitsmarktes, der sich beispielsweise aus technologischen Innovationen und der Globalisierung der Wirtschaftssysteme ergibt, steigt dabei die Nachfrage nach der Vermittlung allgemeiner Fähigkeiten und Schlüsselkompetenzen. Daher stehen neben beruflichen Fähigkeiten und Kenntnissen auch Mathematik- oder Lesekompetenzen immer mehr im Fokus beruflicher Bildungsprogramme. Inwieweit in Deutschland duale oder schulische Ausbildungen allerdings geeignet sind domänenspezifische Kompetenzen zu fördern und ob es Unterschiede zwischen den beiden Systemen gibt, ist bisher empirisch unklar. Befunde zur Kompetenzentwicklung in allgemeinbildenden Kontexten (z. B. Freund et al., 2021a,b) werden bisher nur durch wenige Studien ergänzt, die die Auswirkungen der beruflichen Bildung auf die Mathematik- und Lesekompetenzen (z. B. Brunello & Rocco, 2017; Pensiero & Green, 2018; Nickolaus et al., 2008; Ziegler & Gschwendtner,

2010) untersuchen. Zudem fehlen empirische Arbeiten, die beide beruflichen Ausbildungssysteme innerhalb Deutschlands in den Blick nehmen. Diese Studie fokussiert daher folgende Fragen: Inwieweit entwickeln sich Mathematik- und Lesekompetenzen in unterschiedlichen beruflichen Ausbildungssystemen und welche Unterschiede können in den Systemen beobachtet werden?

Theoretisch ist anzunehmen, dass sich die Kompetenzentwicklung zwischen dualen und schulischen Ausbildungsgängen aufgrund der unterschiedlichen Schwerpunktsetzungen in der Art der Vermittlung von Fähigkeiten und Wissen, den fachlichen Inhalten sowie aufgrund unterschiedlicher didaktischer Herausforderungen unterscheidet. Während beispielsweise die Berufsausbildung im dualen System die betriebliche mit der schulischen Ausbildung kombiniert und ein hoher Anteil der Ausbildungszeit im Betrieb stattfindet, werden fachliche Inhalte in der schulischen Berufsausbildung überwiegend in schulischen Einrichtungen vermittelt. Daher sind die Inhalte des schulischen Teils der dualen Berufsausbildung besonders eng getaktet, so dass im Unterricht nur wenig Zeit bleibt, um neben den beruflichen Inhalten auch allgemeine Kompetenzen zu vermitteln (Roche & Terassi-Haufe, 2016). Dagegen steht in schulischen Ausbildungsgängen mehr Zeit zur Verfügung die Mathematik- und Lesekompetenz im Unterricht zu fördern. Zudem kann die unterschiedliche Zusammensetzung der Schüler*innen in der dualen und schulischen Ausbildung Einfluss auf die Kompetenzentwicklung nehmen, indem beispielsweise aufgrund einer großen Heterogenität unterschiedliche Kompetenzniveaus berücksichtigt werden müssen und dadurch Anforderungen sowie Befürchtungen der Lehrkräfte steigen, Schüler*innen nicht gleichermaßen fördern zu können (Gamoran, 2004). Beispielsweise sind Auszubildende in der dualen Berufsausbildung weitgehend heterogen in Bezug auf den Schulabschluss (Bildungsbericht, 2018), wohingegen in vollqualifizierenden schulischen Berufsausbildungsgängen und berufsvorbereitenden Bildungsgängen die Zusammensetzung der Schülerschaft homogener ist. Wir nehmen daher an, dass in berufsbildenden Klassen der dualen Ausbildung die Lehrkräfte die Mathematik- und Lesekompetenzen nicht in gleichem Maße fördern können und der Kompetenzzuwachs im dualen Ausbildungssegment geringer ausfällt als im schulischen Ausbildungssystem.

Auf Basis der längsschnittlichen Ausbildungsverläufe in der Startkohorte 4 (Klasse 9) des Nationalen Bildungspanels (NEPS) wird die Kompetenzentwicklung in Mathematik und im Lesen in unterschiedlichen Ausbildungskontexten verglichen. Die Analysen konzentrieren sich dabei auf Personen aus dem Haupt- und Realschulzweig, die spätestens nach der 10. Klasse die Schule verlassen haben. Um einen Vergleich nach Ausbildungsformen zu ermöglichen, werden Methoden der Sequenzdatenanalyse angewandt, um Cluster von Ausbildungsverläufen zu generieren und so Ausbildungsverläufe mit Schwerpunkten im dualen und schulischen System zu extrahieren (vgl. Imai & van Dyk 2004, Lopez & Gutman 2017). Zusätzlich wird ein Propensity Score Matching Verfahren eingesetzt, um die Selektion in die Ausbildungsgänge nach verschiedenen Kompetenzniveaus in den Analysen zu berücksichtigen. Dies ermöglicht eine Harmonisierung der Vergleichsgruppen nach individuellen Merkmalen, wie Geschlecht, Kompetenzniveau bei Verlassen der Schule, Schulzweig, familiärem Hintergrund und Migrationshintergrund. Abschließend werden die Kompetenzverläufe auf Basis der gematchten Stichprobe zwischen den Ausbildungsclustern verglichen und die Ergebnisse mit Blick auf praktische Implikationen diskutiert.

Die Rolle der betrieblichen Ausbildungsqualität für die Entwicklung domänenverbundener und domänenspezifischer Kompetenz

Maximilian Krötz[1], Viola Deutscher[1], Beifang Ma[2], Esther Winther[2]
[1]Universität Mannheim, Deutschland; [2]Universität Duisburg-Essen, Deutschland

Trotz der Bedeutung des Arbeitsplatzes als zentralen Lernort der beruflichen Ausbildung sowie des Output-Ziels der beruflichen Handlungskompetenz, liegen bisweilen nur vereinzelt Studien vor, welche die Rolle der betrieblichen Ausbildungsqualität für die Kompetenzentwicklung von Auszubildenden untersuchen (z. B. Baethge-Kinsky, Baethge & Lischewski 2016; Dietzen et al. 2014; Nickolaus, Gschwendtner & Geißel 2009; Pineda-Herrero et al. 2015). Die insgesamt inkonsistenten Befunde scheinen dabei einerseits von der jeweiligen beruflichen Domäne, andererseits aber auch vom Messansatz des Kompetenzkonstrukts sowie von den jeweils gewählten Qualitätsaspekten der Ausbildung abzuhängen. Um der Frage nachzugehen, welchen Effekt die betriebliche Ausbildungsqualität für die Kompetenzentwicklung hat, wurde in dieser Studie ein möglichst breiter Ansatz mit 14 Qualitätsfaktoren sowie Items zur Motivation, mathematischen Interessen und Personenmerkmalen umgesetzt.

Die berufliche Kompetenzentwicklung wurde dabei anhand eines validierten Kompetenztests (AUTOREN) auf zwei Dimensionen erfasst: domänenverbundene und domänenspezifische Kompetenz (Gelman & Greeno 1989). Domänenverbunde Kompetenz umfasst dabei grundlegende Kenntnisse und Fähigkeiten wie etwa allgemeine Rechen- und Lesefähigkeiten, die für das Lösen kaufmännischer Aufgaben benötigt werden (AUTOREN). Domänenspezifische Kompetenz meint Kenntnisse, Regeln und Richtlinien, die berufsspezifisch sind und im Laufe der Ausbildung erlernt werden (AUTOREN). Als Grundlage zur Analyse der betrieblichen Ausbildungsqualität dient das Qualitätsmodell von AUTOREN (2019), welches auf Tynjälä's (2013) „3-P model of workplace learning" aufbaut und die Dimensionen Input, Process und Output umfasst (siehe Fig. 1). Das entsprechende Fragebogeninstrument „VET-LQI" (AUTOREN 2021) wurde zur Erhebung der Ausbildungsqualität verwendet. Anhand einer Kombination von Strukturgleichungsmodellen (SGM) mit latenter Wachstumsmodellierung konnten Längsschnittdaten aus drei Erhebungen von 458 Industriekaufleuten analysiert werden. Aufgrund unterschiedlicher Wachstumsdynamiken sowie der beschriebenen inhaltlichen Differenzierung wurden beide Kompetenzdimensionen separat, mit jeweils eigenen Intercept und Slope Paramatern modelliert (siehe Fig. 2). In einem Dreischritt wurden jeweils zunächst nur Hintergrundmerkmale (Geschlecht, Bildung, Motivation, mathematisches Interesse, etc.) im SGM getestet, bevor ausschließlich Kriterien der Ausbildungsqualität berücksichtigt wurden. Im dritten Schritt wurde das vollständige Modell analysiert (siehe Fig. 3).

Die Ergebnisse zeigen, dass die betriebliche Ausbildungsqualität weniger spezifisch im Sinne einzelner Aspekte[1], sondern eher als holistisches Konstrukt einen deutlichen Einfluss auf die Entwicklung beider Kompetenzdimensionen ausübt. Allerdings scheint die betriebliche Ausbildungsqualität stärker mit domänenverbundener Kompetenzentwicklung (43% Varianzaufklärung) in Zusammenhang zu stehen als mit domänenspezifischer Kompetenzentwicklung (18% Varianzaufklärung). Dies könnte im Wesentlichen auf die Überschneidung der curricularen Inhalte des beruflichen Schulunterrichts mit domänenspezifischen Anforderungen im Testinstrument

zurückzuführen sein, wohingegen die betriebliche Ausbildung einen größeren Einfluss auf allgemeinere Kenntnisse und grundlegende Fähigkeiten auszuüben scheint.

Darüber hinaus zeigt sich, dass persönliche Hintergrundmerkmale (Bildung, Motivation, mathematisches Interesse) vor allem den individuellen Startpunkt zur Kompetenzentwicklung beeinflussen (insbesondere hinsichtlich domänenspezifischer Kompetenz), jedoch weniger für Kompetenzzuwächse im Ausbildungsverlauf verantwortlich sind, da durchweg keine signifikanten Kovarianzen zwischen den Eingangswerten und den Zuwächsen ersichtlich sind.

Hinsichtlich der Ausbildungspraxis legen die Ergebnisse nahe, dass Ausbildungsbetriebe auf eine möglichst breite Herangehensweise bezüglich der Implementierung und des Monitorings von Ausbildungsqualität abzielen sollten. Im Hinblick auf dringend benötigte internationale Vergleichsstudien für den Bereich der beruflichen Ausbildung unterstrichen die Befunde die Bedeutung, berufliche Kompetenz als mehrdimensionales Konstrukt von domänenverbundener und domänenspezifischer Kompetenz zu modellieren.

[1] Als etwas stärker bedeutsam für die Kompetenzentwicklung erwiesen sich lediglich die Methodenvielfalt und Betreuung am Ausbildungsplatz.

Schwierigkeitsbestimmende Aufgabenmerkmale bei Prüfungsaufgaben im gewerblich-technischen Bereich

Pia Schäfer, Elmar Dammann, Felix Walker
Universität Hamburg, Germany

Prüfungsaufgaben sind im gewerblich-technischen Bereich hinsichtlich schwierigkeitsbestimmender Aufgabenmerkmale weitgehend unerforscht. Dabei ist die differenzierte, genaue und valide Erfassung von Kenntnissen und Fähigkeiten, gerade bei Prüfungsaufgaben, sehr bedeutsam, um große Leistungsspektren der geprüften Schüler/innen erfassen zu können. Im Projekt TechKom (Technologiebasierte Kompetenzmessung und -förderung in der Erstausbildung in der Elektro- und Metalltechnik; Förderkennzeichen: 21AP011) werden Aufgabenmerkmale von Prüfungsaufgaben hinsichtlich benötigter Lesekompetenzen untersucht. Studien zeigen, dass gerade bei Berufsschüler/innen Lesekompetenzen schwach ausgeprägt sind (z.B. Lehmann & Seeber, 2007) und dadurch starken Einfluss auf die Prüfungsergebnisse haben können.

In der Untersuchung werden die Bild-Text-Integration (z.B. Mayer & Moreno, 2007; Schnotz & Bannert, 2003) und die Textverständlichkeit (z.B. Brünken et al., 2005; Rabe & Mikelskis, 2007) untersucht. Beide Bereiche sind bisher für den allgemeinbildenden Bereich betrachtet worden (z.B. Ropohl, Walpuski & Sumfleth, 2015). Für die gewerblich-technische Berufsbildung liegen wenige Quellen vor (z.B. Kühn, 2016). Unter Textverständlichkeit werden die Fachsprache, die Textlänge, der Lesbarkeitsindex, die Kohärenz und Hervorhebungen verstanden (Friedrich, 2017; Stawitz, 2010). Bzgl. der Bild-Text-Integration werden der split-attention-Effekt (z.B. Brünken & Leutner, 2001), der picture-textsequencing-Effekt (z.B. Schnotz, 2005) und der rendundancy-Effekt (z.B. Chandler & Sweller, 1991) betrachtet (im Überblick Schnotz, 1994; Schnotz & Bannert, 2003; Mayer, 2001).

Im Projekt werden Aufgaben in den theoretischen Abschlussprüfungen dreier Ausbildungsberufe der gewerblich-technischen Berufsbildung (Mechatroniker/in, Konstruktionsmechaniker/in, Elektroniker/in

für Automatisierungstechnik) untersucht. Dazu wurden Prüfungsaufgaben der PAL (Prüfungsaufgaben und Lehrmittelentwicklungsstelle der IHK Region Stuttgart) aus den letzten 5 Jahren hinsichtlich schwierigkeitsbestimmender Merkmale analysiert (z.B. Lesbarkeitsindex (LIX), Fachbegriffe, Bild-Text-Integration, Split-Attention-Effekt). Für jede Aufgabe wurde ein normalisierter Wert berechnet und mit der Lösungsquote der PAL-Prüfungsaufgaben verglichen. Hinsichtlich der fokussierten Merkmale variierbare Aufgaben wurden anschließend theoriebasiert modifiziert. Im Rahmen einer Interventionsstudie sollen die folgenden Forschungsfragen beantwortet werden: FF3.1: Welchen Einfluss hat die Textverständlichkeit auf die Aufgabenschwierigkeit? FF3.2: Welchen Einfluss hat die Bild-Text-Integration auf die Aufgabenschwierigkeit? sowie FF3.3: Welchen Einfluss hat die Kombination aus Textverständlichkeit und Bild-Text-Integration auf die Aufgabenschwierigkeit? Zur Beantwortung dieser Forschungsfragen wurden für jede ausgewählte Prüfungsaufgabe drei Varianten erstellt, welche 1) die Textverständlichkeit, 2) die Bild-Text-Integration und 3) die Kombination aus 1) und 2) betreffen.

Sowohl die originalen Prüfungsaufgaben als auch deren Variation werden Schüler/innen der o. g. Berufe vorgelegt, die kurz vor ihrer Abschlussprüfung stehen. Das eingesetzte Testheftdesign berücksichtigt die Abhängigkeit der Originalaufgaben mit den jeweiligen Variationen und sichert gleichzeitig die Verankerung aller Items. Die 160 verfügbaren Items je Beruf werden auf 16 Testhefte je Beruf mit jeweils 40 Items verteilt. Die anvisierte Stichprobe liegt bei etwa 500 Schüler/innen pro Beruf. Für den Feldzugang kann auf das Netzwerk der PAL innerhalb der Industrie und Handelskammern zurückgegriffen werden. Erhebungen finden entweder in beruflichen Schulen oder in Ausbildungsbetrieben statt. Die Erhebungen werden rein digital über das Lern-Management-System moodle durchgeführt. Sämtliche Aufgaben und Fragebögen sowie die Testleitung und Einwilligungserklärung sind in dieses System integriert und können von den teilnehmenden Schüler/innen individuell bearbeitet werden.

Erste Analysen aus der Pilotierung (N = 50) zeigen Einflüsse der Aufgabenmerkmale auf die Itemparameter. Dabei liegen die Effekte der Textverständlichkeit und der Bild-Text-Integration auf kleinem bis mittlerem Niveau. Angenommen wird, dass aufgrund der geringen Stichprobe die Effekte nicht signifikant werden. Durch die größere Stichprobe der Hauptstudie erwarten wir den Nachweis signifikanter Effekte und streben latente Schätzungen unter Verwendung des Linear Logistischen Testmodells (z. B. Moosbrugger & Kelava 2020) an.

Im Vortrag werden Ergebnisse der Voranalysen zu den PAL-Prüfungsaufgaben, das Untersuchungskonzept, die digitale Umsetzung der gesamten Testung sowie erste Ergebnisse der Hauptstudie präsentiert. Initiiert durch die Corona-Pandemie werden die Erhebungen zu dieser Untersuchung vollständig digital durchgeführt. Daher erfolgt zusätzlich die Präsentation der digitalen Testplattform und Erfahrungsberichten der digitalen Erhebungen im Vergleich mit Erhebungen in Präsenz.

C-04: Außerunterrichtliche Angebote

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S06 S02 B06
Chair der Sitzung: Thomas Gfrörer

Die Auswirkungen von naturwissenschaftlichen Freizeitaktivitäten auf die Entwicklung von beruflichen Interessen, Berufswünschen, Kompetenzen, Schulnoten und Fähigkeitsüberzeugungen bei Jugendlichen

Thomas Gfrörer, Gundula Stoll, Sven Rieger, Benjamin Nagengast
Hector-Institut für Empirische Bildungsforschung, Germany

Trotz des hohen gesellschaftlichen Bedarfs und den attraktiven Berufsaussichten, entscheiden sich immer noch zu wenig Schülerinnen und Schüler für eine Karriere im MINT-Bereich (Mathe, Informatik, Naturwissenschaft und Technik; Noonan, 2017). Politische Entscheidungsträger versuchen diesem Mangel mit Interventionen an Schulen, wie extracurricularen Kursen oder spezifischen Unterrichtsmethoden entgegenzuwirken (OECD, 2008). Es wird vermutet, dass neben diesen strukturierten Angeboten, die Beschäftigung mit unstrukturierten Freizeitaktivitäten außerhalb der Schule, wie beispielsweise das Lesen eines naturwissenschaftlichen Buches, das Ansehen einer naturwissenschaftlichen Fernsehsendung oder das Recherchieren im Internet über naturwissenschaftliche Inhalte, einen positiven Einfluss auf die Entwicklung von MINT-relevanten Konstrukten haben kann (Dabney et al., 2012; Dou et al., 2019; Henriksen et al., 2015; Maltese & Tai, 2010). Bisherige Befunde deuten auf einen Zusammenhang zwischen naturwissenschaftlichen Freizeitaktivitäten und motivationalen Konstrukten hin (Dabney et al., 2012; Dou et al., 2019). In den meisten empirischen Studien wurden jedoch potenzielle Störvariablen nicht angemessen berücksichtigt. Dies macht es schwierig, kausale Schlussfolgerungen über die Auswirkungen von naturwissenschaftlichen Freizeitaktivitäten zu treffen. Es kann nicht ausgeschlossen werden, dass alternative Mechanismen, wie etwa Selbstselektionseffekte von Personen, die bereits vor den Tätigkeiten ein großes Interesse an MINT-Aktivitäten besaßen, die Effekte erklären. Das Ausmaß in dem naturwissenschaftliche Freizeitaktivitäten die Entwicklung von MINT-bezogenen Konstrukten fördern, ist somit immer noch unklar (Dabney et al., 2012).

Die vorliegende Studie untersucht den Einfluss naturwissenschaftlicher Freizeitaktivitäten auf die Entwicklung von beruflichen Interessen, Berufswünschen, Kompetenzen, schulischen Leistungen und Fähigkeitsüberzeugungen im Bereich MINT. Es stehen somit besonders Konstrukte im Fokus, die sich in empirischen Untersuchungen immer wieder als wichtige Prädiktoren für MINT-Berufe und MINT-Ausbildungswege gezeigt haben (vgl. Wille et al., 2020).

Die Stude greift auf Daten des Nationalen Bildungspanels (NEPS; Blossfeld et al., 2011) zurück. Die untersuchte Teilstichprobe umfasst N = 2.655 Schülerinnen und Schüler aus Deutschland, die im Längsschnitt über die Klassen 9, 11 und 12 hinweg untersucht wurden. Die Freizeitaktivitäten im Bereich Naturwissenschaften wurden anhand von fünf Items gemessen, welche das Schauen naturwissenschaftlicher Fernsehsendungen, das Lesen oder Ausleihen naturwissenschaftlicher Bücher, das Besuchen von Internetseiten mit naturwissenschaftlichen Inhalten, das Lesen naturwissenschaftlicher Artikel in Zeitschriften sowie das Besuchen einer extracurricularen Aktivität im Bereich Naturwissenschaften, umfassen. Die ordinalskalierten Items wurden zu einer dichotomen Variablen aggregiert, um die Schülerinnen und Schüler in Gruppen mit geringem und hohem Engagement einzuteilen. Um die Robustheit der Ergebnisse zu gewährleisten, wurden die Analysen

zusätzlich mit verschiedenen Operationalisierungen (z.B. Median Split, Extremgruppenvergleiche, Kontinuierliche Modellierung und einzelne Items) durchgeführt. Um die kausalen Effekte naturwissenschaftlicher Freizeitaktivitäten auf die Entwicklung von beruflichen Interessen, Berufswünschen, Kompetenzen, schulischen Leistungen sowie Fähigkeitsüberzeugungen im Bereich MINT—unter der Berücksichtigung gemeinsamer Störfaktoren—schätzen zu können, wurde basierend auf den Empfehlungen von VanderWeele et al. (2020), eine sogenannte „outcome-wide analysis“ durchgeführt. Bei dieser Form der statistischen Analyse, wird der Einfluss eines Treatments auf eine Vielzahl von Konstrukten untersucht.

Die Ergebnisse zeigen, dass naturwissenschaftliche Freizeitaktivitäten einen Einfluss auf investigative berufliche Interessen, nicht aber auf Berufswünsche, Kompetenzen, Schulnoten und Fähigkeitsüberzeugungen haben. Die Ergebnisse lassen darauf schließen, dass Jugendliche mit ähnlichem dispositionellen Interesse, die vermehrt naturwissenschaftliche Freizeitaktivitäten aufsuchen, ein stärkeres investigatives Interesse ausbilden können als Jugendliche, die weniger naturwissenschaftliche Freizeitaktivitäten aufsuchen. Dieses Befundmuster ist im Einklang mit zentralen theoretischen Modellen zur Interessentwicklung (Hidi & Renninger, 2006; Holland, 1997; Su et al., 2019), in denen das wiederholte und kontinuierliche Engagement in bestimmten Aktivitäten jeweils eine Schlüsselkomponente für die Entwicklung von Interessen darstellt. Übertragen auf unserer Studie lässt sich annehmen, dass Schülerinnen und Schüler, die sich in ihrer Freizeit für naturwissenschaftliche Aktivitäten entscheiden und dies wiederholt tun, vermehrt positive Erfahrungen diesbezüglich machen. Diese positiven Erfahrungen könnten wiederum zu einem wiederholten Aufsuchen dieser Tätigkeiten und einer weiteren Vertiefung der Interessen führen. Unsere Studie liefert damit wichtige Erkenntnisse über die Entwicklung von Interessen im MINT-Bereich.

Die Erlebnisqualität außerunterrichtlicher Angebote: Eine quantitative Analyse der Effekte auf die Entwicklung der Mathematik- und Deutschnoten

Till Woller[1,2], Buhl Monika[1], Reschke Katharina[1]
[1]Ruprecht-Karls-Universität Heidelberg; [2]Universität Mannheim

Diese Studie untersucht den Einfluss von Tiefenstrukturen der Schüler:innen bei außerunterrichtlichen Angeboten auf die Entwicklung von Schulnoten anhand des Konzepts der *quality of experience* nach Shernoff (2010), im Folgenden Erlebnisqualität. Shernoffs theoretische Grundannahmen basieren auf der Flow-Theorie von Csikszentmihalyi (1990), der zufolge sich die Lebensqualität eines Individuums steigert, wenn es einen Flow-Zustand erlebt. Aufbauend auf dem Konzept von Schmidt et al. (2007) identifizierte Shernoff drei Bedingung für eine hohe Erlebnisqualität: erstens ein hohes Maß an Herausforderung, um die Ziele einer Aktivität zu erreichen (*challenge*); zweitens ein hoher Einsatz von Fähigkeiten bei der Aktivität (*skills*) und drittens, dass die Aktivität als relevant betrachtet wird (*importance*). Das Flow-Erleben (*engagement*) definierte Shernoff als gleichzeitige Erfahrung von *concentration*, *interest* und *enjoyment*.

Shernoff befand signifikant positive Effekte höherer *importance* und *challenge* Werte auf Schulnoten. Ihm fehlte es jedoch an empirischer Evidenz aus einer repräsentativen Stichprobe. Diese Forschungslücke schließt diese Arbeit. Mit Hilfe eines tiefgreifenden Verständnisses des Zusammenhangs zwischen Erlebnisqualität und Lernerfolg könnte bei der Gestaltung von

Lernarrangements die Erlebnisqualität erhöht werden, um die Leistungsentwicklung der Schüler:innen zu begünstigen.

Diese Arbeit verfolgt das Ziel, zu überprüfen, ob einzelne Faktoren der Erlebnisqualität außerunterrichtlicher Angebote bereits eine Auswirkung auf die Entwicklung von Schulnoten haben oder sich bei der hohen Ausprägung aller Faktoren und dem Eintreten eines Flow-Zustandes eine Verbesserung der Mathematik- und Deutschnoten nachweisen lässt. Die zentrale Forschungsfrage lautet: Wie wirken die einzelnen Faktoren und das Gesamtkonstrukt Erlebnisqualität außerunterrichtlicher Angebote von Shernoff (2010) auf die Entwicklung der Mathematik- und Deutschnoten bei Schüler:innen der Sekundarstufe 1? Auf Basis von Shernoffs Theorie zur Wirkung von Erlebnisqualität auf schulische Outcomes ist zu vermuten, dass sich sowohl die einzelnen Faktoren der Erlebnisqualität als auch das Gesamtkonstrukt positiv auf die Entwicklung der Mathematik- und Deutschnoten der Schüler:innen auswirken.

Die empirischen Analysen wurden anhand des Längsschnittdatensatzes der ersten StEG-Studie vorgenommen, für die *N*=6.853 Schüler:innen an drei Messzeitpunkten im Zeitraum von 2005 bis 2009 befragt worden sind. Zur Identifikation etwaiger signifikanter Effekte von *engagement*, *importance*, *skills*, *challenge* und Erlebnisqualität auf die Entwicklung der Mathematik- und Deutschnoten wurden zehn latente Wachstumskurvenmodelle in R berechnet und zu jedem Modell ein Hypothesentest durchgeführt.

Neun von zehn Hypothesen wurden verworfen. *Engagement* wirkte, wie erwartet, negativ jedoch weder signifikant auf die Mathematik- noch auf die Deutschnoten. Für *importance* wurden keine signifikanten Effekte gefunden. Die Effekte von *skills* auf die Entwicklung der Mathematik- und Deutschnoten waren, entgegen der Erwartung, positiv und zu einem Messzeitpunkt signifikant. Der Faktor *challenge* wirkte in unterschiedliche Richtungen und zu keinem Messzeitpunkt signifikant auf die Mathematiknoten. Auf die Entwicklung der Deutschnoten wirkte *challenge* signifikant negativ. Die entsprechende Hypothese wurde bestätigt. Es wurden keine signifikanten Effekte von Erlebnisqualität auf die Mathematik- oder Deutschnoten gefunden.

Die Faktoren *challenge* und *skills* weisen eine niedrige Inhaltsvalidität auf, ebenso wie das Attribut *concentration*. Sämtliche anderen Faktoren und Attribute wurden valide operationalisiert.

Die Befundlage wirft die Diskussion auf, ob eine hohe Ausprägung des Faktors *skills*, wie angenommen, auf eine bessere Notenentwicklung hindeutet oder ein gegenteiliger Effekt angenommen werden sollte. Falls sich Zweiteres bestätigt, sollte *skills* aus dem theoretischen Konstrukt eliminiert werden. Aufgrund der durchweg niedrigen Bewertung des Faktors *importance*, ist zu bezweifeln, dass die Befragten einen Flow-Zustand während der außerunterrichtlichen Angebote erlebten.

Der Faktor *challenge*, der die Notenentwicklung im Fach Deutsch begünstigte, kann durch eine bessere Binnendifferenzierung in den Ganztagsangeboten gesteigert werden. Er birgt das Potential, sich positiv auf die Leistungsentwicklung von Schüler:innen auszuwirken. Die niedrigen *importance*-Werte offenbaren den Bedarf, die Relevanz der Lerninhalte in außerunterrichtlichen Angeboten sowohl deutlicher hervorzuheben als auch die Auswahl der Inhalte daran zu orientieren, was den

Schüler:innen als wichtig erscheint. Shernoff befand, dass hohe *importance*-Werte die Notenentwicklung signifikant verbessern. Dieses Potential sollte in außerunterrichtlichen Angeboten besser ausgeschöpft werden.

Fördert der Besuch einer Museumsausstellung die Energiekompetenz von Schüler:innen?

Sarah Kellberg[1,2], Doris Lewalter[1], Jeffrey Nordine[3], Melanie Keller[2], Stephanie Moser[1]
[1]TUM School of Social Sciences and Technology, Technische Universität München; [2]Didaktik der Physik, IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik; [3]College of Education, University of Iowa

Ziel von Schule ist es u.a. Schüler:innen an die großen Herausforderungen unserer Zeit, wie beispielsweise Klimawandel oder Energieversorgung, heranzuführen und eine Auseinandersetzung mit diesen Themen anzuregen. Dafür könnte die Kombination aus Schulunterricht und modernen Ausstellungen zu sozial-/naturwissenschaftlichen Themen einen ganzheitlichen und zeitsparenden Ansatz darstellen (Dunlop et al., 2018).

Allerdings ist derzeit noch nicht bekannt, ob moderne Ausstellungen, die sich mit kritischen Themen auseinandersetzen und ihren Besucher:innen ein hohes Maß an Aktivität und Interaktion anbieten, (1) tatsächlich die in sie gesetzten Erwartungen erfüllen und (2) inwieweit sich schulisches Wissen, wie es derzeit gelehrt wird, und inhaltliches Interesse als nützlich für das Lernen in solchen Ausstellungen erweist.

Die vorliegende Studie untersucht diese Fragestellungen anhand der drängenden sozial-/naturwissenschaftlichen Herausforderung der Energiewende, für deren Gelingen eine kritische Energiekompetenz (energy literacy) der Gesellschaft notwendig erscheint. Wir untersuchten, inwieweit der Besuch einer sozial-/naturwissenschaftlichen Ausstellung zur Energiewende die Energiekompetenz von Schüler:innen fördert, und inwieweit die potenziellen Effekte vom thematischen Interesse der Schüler:innen und ihrem, in der Schule erworbenen, konzeptionellen Energiewissen beeinflusst werden.

Im Rahmen einer Fragebogenstudie mit Prä-Post-Design haben Schüler:innen (N= 185; 8. bis 10. Jahrgangsstufe, 42% Mädchen) in einem naturwissenschaftlichen Museum eine Ausstellung zum Thema Energiewende besucht. Die Ausstellung konzentrierte sich auf den globalen Prozess der Energiewende und seine sozialwissenschaftlichen Aspekte und betonte die Möglichkeiten der Öffentlichkeit, kreativ zum Prozess der Energiewende beizutragen. Sie richtete sich in erster Linie an Jugendliche und junge Erwachsene und wies ein innovatives Format mit einem hohen Maß an Interaktivität auf: Besucher:innen der Ausstellung wurden darin eingeladen in die Rolle eines Politikers/einer Politikerin zu schlüpfen und mit Hilfe von Informationen aus der Ausstellung und in Auseinandersetzung mit verschiedenen Akteuren der Energiewende sowie potentiellen Wähler:innen sich für bestimmte energiepolitische Maßnahmen zu entscheiden und so letztlich ihre persönliche Energiewende zu gestalten. Im Rahmen der Studie wurde ein mixed-methods Ansatz zur Erfassung der Energiekompetenz gewählt. Dazu wurden eine offene Frage zur Erklärung des Begriffs Energiewende mit einer im Anschluss quantifizierten Abfrage des Argumentrepertoires zu konventionellen und erneuerbaren Energien (in Anlehnung an Knipfler, 2008; Toplak & Stanovich, 2003) verwendet, ebenso wie zwei Skalen zur Bewusstheit der gesellschaftswissenschaftlichen Dimensionen der Energiewende im Sinne der socio-scientific issues („open inquiry“ und „skepticism“

Sadler et al., 2007). Die Abfrage erfolgte eine Woche vor und eine Woche nach dem 90-minütigen Ausstellungsbesuch. Vor dem Ausstellungsbesuch wurde zudem das thematische Interesse der Schüler:innen in Anlehnung an Krapp (2002) und ihr konzeptionelles Wissen über die Energiewende (Neumann et al., 2013) erfasst.

Die Ergebnisse von Wilcoxon-signed rank-Tests weisen auf eine signifikante Förderung der verschiedenen Indikatoren der Energiekompetenz der Schüler:nnen durch den Ausstellungsbesuch, mit niedrigen bis zu mittleren Effekten hin. So wurde für das Wissen zum Begriff Energiewende, das Argumentrepertoire sowie die Offenheit der Thematik (open inquiry) kleine positive Effekte und für die Skepsis gegenüber ggf. verzerrten Informationen (skepticism) ein mittlerer positiver Effekt ermittelt. Allerdings konnten wir mithilfe multipler linearer Regressionen keinen signifikanten Einfluss des thematischen Interesses sowie des konzeptionellen Energiewissens auf dieses Ergebnis feststellen. Zusammenfassend kann die Annahme, dass der Ausstellungsbesuch einen positiven Effekt auf die Energiekompetenz hat in Teilen bestätigt werden. Demgegenüber kann die Annahme, dass das in der Schule erworbene konzeptuelle Energiewissen als auch das inhaltliche Interesse für das Lernen in diesem informellen Kontext förderlich ist, nicht bestätigt werden. Die Befunde werden sowohl in Hinblick auf die Grenzen der Wirksamkeit schulischen Wissens für außerschulisches Lernen als auch bezogen auf die Potentiale außerschulischer Lernumgebungen diskutiert.

Unterstützung des selbstgesteuerten Lernens von Schülerinnen und Schülern im Museum durch motivationale und metakognitive Prompts

Katrin Neubauer, Doris Lewalter
TUM School of Social Sciences and Technology, Department Educational Sciences, Professur für Formelles und Informelles Lernen, München, Deutschland

Ziel schulischer Bildung ist es, u.a. das Fundament für lebenslanges Lernen sowie aktive Teilhabe an gesellschaftlichen Prozessen und die Mitgestaltung der eigenen Lebensumwelt zu legen. Hierfür stellt die Fähigkeit zur selbständigen Wissensaneignung nicht nur im schulischen Kontext, sondern auch in informellen Lernkontexten, wie Museen, eine wichtige Voraussetzung dar (Düx & Sass, 2005; Hungerland & Overwien, 2004).

Gerade Museen eignen sich aufgrund ihrer situativen Rahmenbedingungen und der Möglichkeit zur freien Erkundung (z.B. Behrendt & Franklin, 2014; Whitesell, 2016) dafür Lernerfahrungen zum selbstgesteuerten Lernen am informellen Lernort zu sammeln. Studien zeigen allerdings, dass es häufig wenig effektiv ist, wenn sich Schüler:innen Inhalte im Museum komplett eigenständig erarbeiten (z.B. Bamberger & Tal, 2007; Sung et al., 2008) und dass zudem für das Erreichen positiver motivationaler und kognitiver Lerneffekte ein minimales Maß an Strukturierung und Unterstützung notwendig ist (z.B. Neubauer, 2015). In anderen Lernkontexten haben sich instruktionale Prompts als effektive Unterstützungsmaßnahme selbstgesteuerten Lernens erwiesen (zsfd. Bannert, 2009). Im Museumskontext liegt hierzu jedoch kaum Forschung vor (z.B. Yoon et al., 2012; 2013, 2018).

Hier setzt das Projekt „ProMus“ an und untersucht, welche Art instruktionaler Prompts die Schüler:innen während ihres selbstgesteuerten Lernens im Museum zu einer effektiveren und

aktiveren Beschäftigung mit den Exponaten anregen kann. Dazu werden folgende Fragestellungen im Kontext schulischer Museumsbesuche untersucht:

1. Inwieweit und wie werden verschiedene Promptarten von den Lernenden genutzt?
2. Welches Engagement-Level zeigen die Lernenden in Abhängigkeit verschiedener Promptarten?
3. Welchen Einfluss haben verschiedene Promptarten auf motivationale und kognitive Lernergebnisse der Lernenden?

Zur Klärung dieser Fragestellungen wurden bisher 135 Schüler:innen [59,3 % weiblich; M_{Alter} = 15.80 (SD = .87)] der 10. Jahrgangsstufe Gymnasium im Rahmen einer quasiexperimentellen Studie im Pretest-Posttest-Design vor und direkt nach ihrem Besuch zum Thema „Mobilität und Mobilitätswende“ im Verkehrszentrum des Deutschen Museums München schriftlich befragt (Abschluss Datenerhebung Ende Oktober 2022). Die Schüler:innen jeder Klasse wurden anhand von Kleingruppen zufällig auf die drei Treatmentversionen [*Promptart:* (1) metakognitiv, (2) motivational, (3) metakognitiv + motivational] und (4) Kontrollgruppe (keine Prompts) aufgeteilt (n = 31-36).

Zur Erfassung des thematischen (Vor-)Wissens wird eine selbstentwickelte Ratingskala und ein Wissenstest, der Faktenwissen und Verständnis entsprechend der Bloom´schen Taxonomie (1972) erfasst, verwendet. Das thematische Interesse wird in Anlehnung an Krapp (2002, 2006) und die besuchsbezogene selbstbestimmte Lernmotivation mit einer Skala von Frey et al. (2009) erhoben. Das Engagement-Level der Lernenden (Interaktion mit Exponaten und Mitlernenden) wird mithilfe des dreidimensionalen Activity-Engagement-Survey von Ben-Eliyahu et al. (2018) und die individuelle Nutzungshäufigkeit und –form der Prompts mit vier selbstentwickelten Items erfasst (Manipulation check).

Erste varianzanalytische Ergebnisse der bereits vorliegenden Stichprobe zeigen, dass alle Treatmentgruppen ihre entsprechenden Prompts „meistens/immer“ genutzt haben, wobei die Nutzungsintensität bei den motivationalen Prompts am höchsten ausfällt. Alle drei Promptarten wurden von den Lernenden meist als Anstoß gesehen, in welche Richtung sie denken, handeln und diskutieren sollen.

Alle Gruppen geben nach dem Besuch ein mittleres thematisches Interesse an, wobei der Wert für die Kontrollgruppe am niedrigsten ausfällt. Die besuchsbezogene selbstbestimmte Lernmotivation der Kontrollgruppe und der Mot-Gruppe ist tendenziell höher als die der anderen beiden Gruppen. Alle Vergleichsgruppen berichten ein hohes Engagement-Level, wobei der Wert der Metakog-Gruppe tendenziell signifikant höher ausfällt. Bei allen Untersuchungsgruppen zeigt sich ein signifikanter Wissenszuwachs, welcher bei der Kombi-Gruppe im Vergleich zur Mot- und Kontrollgruppe tendenziell signifikant höher ausfällt.

Die Ergebnisse deuten an, dass die Gruppe ohne Prompt geringere Ausprägungen in den Lernergebnisvariablen und im Engagement-Level zeigt als die Prompt-Gruppen, wobei bei diesen die Höhe der Ausprägungen von der Passung zwischen Prompt-Ausrichtung und Ausrichtung der Variablen abhängt. Finale Ergebnisse werden im Rahmen der GEBF-Tagung präsentiert und hinsichtlich ihrer Bedeutung für das Verständnis und die Förderung kognitiver und motivationaler Merkmale selbstgesteuerter Lernprozesse von Schüler:innen in informellen Kontexten diskutiert.

C-05: Unterrichtsqualität

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S06 S03 B06
Chair der Sitzung: Alfred Lindl

Modellierung fachspezifischer Merkmale von Unterrichtsqualität in den sechs Unterrichtsfächern Deutsch, Englisch, Ev. Religion, Latein, Mathematik und Musik

Alfred Lindl, Stefan Böhringer, Patrick Ehrich, Maria Gürtner, Maximilian Gutsmiedl, Annika Krämer, Mara Rader, Laura Simböck
Universität Regensburg, Deutschland

Merkmale qualitätsvollen Unterrichts sowie deren Zusammenhänge mit unterrichtlichen Zielkriterien stehen nicht erst in den letzten Jahren im Fokus der empirischen Bildungsforschung, wie die zahlreichen Themenhefte hierzu vermuten ließen, die kürzlich in einschlägigen wissenschaftlichen Zeitschriften erschienen sind. Während zunächst normative Kriterienlisten für guten Unterricht vorgelegt wurden, die generische Gültigkeit beanspruchten (z. B. Brophy, 2000; Helmke, 2021; Meyer, 2004), setzt der gegenwärtige Forschungsdiskurs andere Schwerpunkte: Einerseits rückt gemäß Berliner (2005) das Kriterium evidenzbasierter Effektivität in den Vordergrund (z.B. Praetorius et al., 2018; Praetorius et al., 2022), andererseits aufgrund der eminenten Bedeutung fachbezogener Unterrichtsprozesse für den Lernerfolg von Schüler:innen (Seidel & Shavelson, 2007) fachspezifische Merkmale von Unterrichtsqualität. Unsicherheiten und Desiderate in der Lehrkräftebildung und Bildungswissenschaft liegen also bezüglich der empirisch überprüften Evidenz bisheriger Annahmen zu unterrichtlichen Qualitätsmerkmalen und des deutlichen Forschungsbias zugunsten mathematisch-naturwissenschaftlicher Fächer vor (Lindmeier & Heinze, 2020; Praetorius et al., 2018, 2020b).

Dies greift der vorliegende Beitrag der interdisziplinären Forschungsgruppe FALKO-PV (Fachspezifische Lehrkraftkompetenzen – Prädiktive Validierung) auf und nimmt in zwei aufeinander bezogenen Untersuchungsansätzen, deren erste Ergebnisse präsentiert werden sollen, die Unterrichtsfächer Deutsch, Englisch, Evangelische Religion, Latein, Mathematik und Musik in den Blick. Leitfragen sind hierbei, inwiefern sich existierende Modelle von Unterrichtsqualität (z.B. Praetorius et al., 2020a) auf die genannten Fächer übertragen lassen, inwiefern diese um spezifische Merkmale aus bestimmten Domänen zu ergänzen sind, und schließlich, wie sich postulierte Merkmale – generisch oder fachspezifisch – konzeptualisieren, operationalisieren und validieren lassen.

Hierzu wurde zunächst in Literaturrecherchen zu allen sechs beteiligten Fächern analysiert, welche Prozessmerkmale von Unterricht in der jeweiligen Fachdidaktik besonders diskutiert werden, um potenzielle fachspezifische Qualitätsaspekte zu identifizieren. Da sich der diesbezügliche Forschungsstand teils als defizitär erwies (vgl. Kranefeld, 2021; Lindl & Krämer, 2022; Rothgangel, 2021; Wilden, 2021), wurden im Frühjahr 2022 pro Fach als Delphi-Studien geplante Onlineumfragen mit Lehramtsstudierenden, Lehrkräften sowie weiteren Personen, die in die Lehrkräftebildung involviert sind, mit einem einheitlichen Design durchgeführt. In diesen sollten die Teilnehmenden (N = 367) zunächst in Freitextfeldern drei für ihr Fach absolut spezifische Qualitätskriterien beschreiben, daraufhin die Relevanz verschiedener vorgegebener und aktuell diskutierter Merkmale in Hinblick auf ihr Unterrichtsfach auf einer fünfstufigen Likertskala bewerten. Die offenen Antworten wurden jeweils von (mind.) zwei fachkundigen Kodierer:innen unabhängig voneinander gemäß einem deduktiven Beurteilungsraster, das sich am Syntheseframework von Praetorius et al. (2020a) orientiert, kategorisiert, wobei die Interraterreliabilität in jedem Fach angemessen war (78 % ≤ proz.

Übereinstimmung ≤ 92 %); die Relevanzeinschätzungen wurden itemweise und nach inhaltlichen Dimensionen aggregiert ausgewertet. Die Analysen zu den offenen Items lassen in allen Fächern einen klaren Fokus auf Auswahl und Thematisierung von Fachinhalten und -methoden erkennen, wohingegen Merkmale wie kognitive Aktivierung in Englisch oder formatives Assessment in Latein keine Erwähnung finden. Dafür treten fachspezifische Merkmale wie Intermedialität in Deutsch oder Subjektorientierung in Religion hervor. In Varianzanalysen zu den mittleren Relevanzeinschätzungen werden zwischen den Fächern Unterschiede in Hinblick auf alle sieben inhaltlichen Dimensionen signifikant und fachbezogene Schwerpunkte deutlich (z.B. vergleichsweise hohe Werte für kognitive Aktivierung in Deutsch, niedrige für sozio-emotionale Unterstützung in Latein).

Auf den Ergebnissen dieser Befragungen fußt ein digitaler Fragebogen zur Beurteilung von Unterrichtsqualität aus der Perspektive von Schüler:innen, der neben generischen einen besonderen Fokus auf fachspezifische Merkmale legt. Hierzu wurden für die generischen Kriterien mehr als 1000 Items aus bestehenden Instrumenten gesichtet und selektiert sowie für die fachspezifischen Merkmale eigenständig neue Items konstruiert. Die aktuelle Version des Fragebogens umfasst (fachabhängig etwa) 180 Items, die auf fünfstufigen Likertskalen zu beantworten sind und sich neun Dimensionen zuordnen lassen. In einer ersten Pilotierung mit ca. 1000 Schüler:innen aller Fächer weist er bezüglich aller Skalenreliabilitäten gute interne Konsistenzen auf (Cronbachs Alpha ≥ .83). Diese und weitere Ergebnisse der noch nicht abgeschlossenen Pilotierung sollen im Vortrag ebenfalls vorgestellt und diskutiert werden.

„Es ist unmöglich, diesen unruhigen Haufen zu unterrichten!“ – Zusammenhänge von Unterrichtsqualität und Klassenkomposition in Klassen derselben Lehrkraft

Anne Heinschel, Sofie Henschel, Camilla Rjosk
Institut zur Qualitätsentwicklung im Bildungswesen

Theorie und Fragestellungen

In der Wahrnehmung vieler Lehrkräfte gibt es einfach zu unterrichtende Klassen und solche, in denen eine hohe Unterrichtsqualität (kognitive Aktivierung, Klassenmanagement, konstruktive Unterstützung; vgl. Klieme et al., 2009) kaum zu erreichen ist (Tiedemann & Billmann-Mahecha, 2002). Nach dem Angebots-Nutzungs-Modell (Helmke, 2017) entsteht die Unterrichtsqualität durch das Angebot der Lehrkraft und dessen Nutzung durch die Schüler:innen. Insofern können Merkmale der Lehrkraft und der Klasse für die Unterrichtsqualität bedeutsam sein (Bell et al., 2012; Palardy, 2013).

Bislang wird die Qualität des Unterrichts häufig primär der Lehrkraft zugeschrieben (u.a. Baumert et al., 2010; Kunter et al., 2013). Ohne Unterscheidung von Lehrkraft- und Klassenebene zeigen einige Studien allerdings, dass Klassenmerkmale die Ausprägung von Unterrichtsqualität mitbedingen: insbesondere proximale, also durch Unterricht veränderbare Faktoren der Klasse (kognitive und motivationale Merkmale wie Leistung, Selbstkonzept und Interesse) und weniger distale Faktoren (z.B. Geschlecht, Zuwanderungshintergrund, Trautmann & Wischer, 2011) hängen mit der Unterrichtsqualität zusammen (Becker & Keller, 2022; Schiepe-Tiska et al., 2016). Sofern Dimensionen der Unterrichtsqualität primär durch die Klasse bestimmt werden, wäre anzunehmen, dass diese stärker zwischen Klassen als zwischen Lehrkräften variieren (Fauth et al., 2020; Voss et al., 2022). Dies gilt insbesondere für das Klassenmanagement (Lei et al., 2018; Kane & Staiger,

2012), da hier meist explizit das Schülerverhalten betrachtet wird (Fauth et al., 2020). Fauth et al. (2020) untersuchten die Unterrichtsqualität einer einzelnen Lehrkraft und zeigten, dass insbesondere proximale Faktoren für Qualitätsunterschiede zwischen Klassen bedeutsam waren. Bislang ist unklar, ob dieses Muster über Lehrkräfte hinweg zu beobachten ist. Vor diesem Hintergrund untersuchen wir folgende Fragestellungen:

1) Weist Unterrichtsqualität höhere Varianzanteile auf Klassen- oder auf Lehrkraftebene auf?

2) Inwiefern hängen Unterschiede in der Unterrichtsqualität in Klassen derselben Lehrkraft mit Merkmalen der Klassenkomposition zusammen?

Methode

Es werden Daten von 210 nicht-gymnasialen Klassen (2 Klassen pro Lehrkraft, insg. 3493 Schüler:innen) ausgewertet, die im IQB-Bildungstrend 2018 u. a. Fragen zur Qualität (kognitive Aktivierung/KA, 12 Items, α = .81; konstruktive Unterstützung/KU, 11 Items, α = .90; Klassenmanagement/KM, 7 Items, α = .79) ihres Mathematikunterrichts beantworteten (Stanat et al., 2019). Außerdem werden distale (Geschlecht; sozioökonomischer Hintergrund: HISEI; sprachlicher Hintergrund: ein- vs. mehrsprachig; Note) und proximale Merkmale (Mathe-Selbstkonzept: 4 Items, α = 0.88; Interesse: 4 Items, α = 0.89) berücksichtigt. Zur Prüfung von Fragestellung 1 werden mit Mehrebenenanalysen Varianzanteile der Unterrichtsqualität auf Ebene der Schüler:in (L1), Klasse (L2) und Lehrkraft (L3) bestimmt. Zur Prüfung von Fragestellung 2 gehen pro Klassenpaar auf L2 Differenzwerte in eine Regressionsanalyse ein (AV: Unterrichtsqualität, UV: Klassenkompositionsmerkmale). Die Koeffizienten bilden ab, inwiefern eine höhere (niedrigere) Unterrichtsqualität mit einer höheren (geringeren) Ausprägung der Kompositionsmerkmale in der einen oder anderen Klasse einer Lehrkraft einhergeht.

Ergebnisse und Diskussion

Die Ergebnisse zeigen, dass die Dimensionen der Unterrichtsqualität auf Klassen- und Lehrkraftebene ähnlich hohe Varianzanteile aufweisen, sich das Verhältnis von Klasse zu Lehrkraft jedoch leicht unterscheidet (KA: Klassenanteil höher, KU: etwa gleich, KM: Lehrkraftanteil höher). Somit scheint der Handlungsspielraum für Lehrkräfte insbesondere im Klassenmanagement relativ unabhängig von der unterrichteten Klasse zu sein, wohingegen er in der kognitiven Aktivierung durch Eigenschaften der Klasse begrenzt scheint. Die Ergebnisse zu Fragestellung 2 zeigen, dass distale Faktoren überwiegend nicht mit der Unterrichtsqualität zusammenhängen. Einzig ein höherer Anteil von mehrsprachigen Schüler:innen hängt mit einer geringeren KA zusammen. Lehrkräfte passen das kognitive Anforderungsniveau ihres Unterrichts offenbar je nach sprachbezogener Klassenzusammensetzung an. Unter den proximalen Faktoren hängt das mittlere fachbezogene Interesse positiv mit der konstruktiven Unterstützung zusammen. Weist eine von zwei Klassen derselben Lehrkraft ein höheres Interesse auf, fühlt sich diese Klasse besser unterstützt. Insgesamt scheint Unterricht zwar im Zusammenspiel von Lehrkraft und Klasse stattzufinden, Merkmale der Klasse allein erklären aber nur sehr bedingt Unterschiede in der Unterrichtsqualität derselben Lehrkraft.

Der Einfluss von Lehrkräftemerkmalen auf die Unterrichtsqualität: Eine Metaanalyse

Nicoletta Bürger, Katharina Engelmann, Laura Schultze, Gabriele Prinz, Marvin Scott Schlamelcher, Jürgen Menthe, Barbara Schmidt-Thieme, Christof Wecker
Stiftung Universität Hildesheim, Deutschland

Die Qualität von Unterricht gilt als besonders relevant für die Bildung junger Menschen, da bereits vielfach in empirischen Studien nachgewiesen wurde, dass Unterrichtsqualität - durch Lehrkräftemerkmale beeinflusst - Auswirkungen auf den Lernzuwachs von Schülerinnen und Schüler hat (Blömeke et al., 2022; Klusmann et al., 2022). Das im deutschsprachigen Raum etablierte Verständnis von Unterrichtsqualität, wie es beispielsweise von Klieme et al. (2009) beschrieben wird, beinhaltet die Subdimensionen Lernunterstützung, Klassenführung und kognitive Aktivierung. Zu den Einflussfaktoren auf die Unterrichtsqualität gehören nach Baumert und Kunter (2006) Lehrkräftemerkmale wie das professionelle Wissen der Lehrkräfte (d. h. Fachwissen, fachdidaktisches Wissen und pädagogisch-psychologisches Wissen) sowie Überzeugungen und die motivationale Orientierung, deren Zusammenhang mit der Unterrichtsqualität empirisch belegt wurde (Kunter, 2011; Voss et al., 2011).

In der Literatur fehlt derzeit eine Zusammenfassung der Lehrkräftemerkmale und eine kritische Bewertung ihrer relativen Bedeutung für die Unterrichtsqualität. Eine solche Zusammenfassung und Bewertung ist jedoch für die Lehrerbildung für die Förderung spezifischer Lehrkräftemerkmalen wichtig. Mit einer wachsenden Zahl empirischer Studien zur Unterrichtsqualität steigt die Notwendigkeit, die verfügbare empirische Evidenz systematisch zu bündeln. Diese Arbeit untersucht daher die folgende Fragestellung: Inwieweit gibt es einen Zusammenhang zwischen Lehrkräftemerkmale und Unterrichtsqualität?

Um diese Fragestellung zu untersuchen, wird eine Metaanalyse durchgeführt. Inklusionskriterien waren (a) die Messung eines Lehrkräftemerkmal, (b) die Messung von Unterrichtsqualität oder eine ihrer Subdimensionen in Bezug auf (c) Unterricht an einer Regelschule (d) in Präsenzunterricht und (e) der Bericht von quantitativen Ergebnissen. Für die Suche geeigneter wissenschaftlicher Befunde wurden Suchbegriffe in Bezug auf die Unterrichtsqualität sowie die Subdimensionen Klassenführung, kognitive Aktivierung und Lernunterstützung (Klieme et al., 2009) und Suchbegriffe in Bezug auf Lehrkräftemerkmale, also Einstellungen, sozialer Einfluss und Einschätzungen der Verhaltenskontrolle, wie Selbstwirksamkeit oder Angst (Baumert und Kunter 2006), in einem booleschen Suchausdruck kombiniert. Die Literaturrecherche wurde in ERIC, PsycINFO und Psyndex durchgeführt. Die Suche wurde auf veröffentlichte Publikationen auf Englisch oder Deutsch eingeschränkt. Die Suchanfrage ergab 7799 Treffer (1. August 2022). Die Abstracts werden zunächst nach den oben genannten Einschlusskriterien gescreent. Aktuell werden vier Personen für die Codierung der Abstracts trainiert. Die Volltexte der inkludierten Studien werden auf Charakteristika der Untersuchungen (z. B. Erfahrung der Lehrkräfte, von den Lehrkräften unterrichtete Klassenstufen und Fächer, Art und Erhebungsinstrument des Unterrichtsqualitätsmaßes) codiert. Aus den berichteten Ergebnissen werden Korrelationskoeffizienten zur Berechnung der Zusammenhänge zwischen den Variablen extrahiert. Die statische Analyse wird mit metaanalytischen Strukturgleichungsmodellen (MASEM) durchgeführt (Jak & Cheung, 2020; Jak et al., 2021).

Ergebnisse erster Teilanalysen zeigen, dass sowohl die Selbstwirksamkeit von Lehrkräften (k = 25, N = 10985, r = 0.167, p < .001) als auch das Professionswissens von Lehrkräften (k = 12, N = 2039, r = 0.143, p = .030) einen signifikanten kleinen positiven Zusammenhang mit Unterrichtsqualität haben. Es hat sich kein signifikanter Zusammenhang zwischen Selbstwirksamkeit und Professionswissen (k = 5, N = 361, r = -0.045, p = .710) gezeigt.

Mit der metaanalytischen Strukturgleichungsmodellierung werden die relativen Zusammenhänge der Lehrkräftemerkmale in Form von Einstellungen, soziale Normen und Verhaltenskontrolle in Bezug auf die Unterrichtsqualität untersucht. Diese Metaanalyse wird, basierenden auf dem aktuellen wissenschaftlichen Stand zu den Zusammenhängen von einzelnen unterschiedlichen Lehrkräftemerkmale mit Unterrichtsqualität, einen neuen Einblick darin bieten, welchen Lehrkräftemerkmale sich als systematische Prädiktoren für Unterrichtsqualität erweisen. Die Ergebnisse können der Lehrkräfteausbildung bereits wichtige Hinweise darauf geben, welche Aspekte besonders relevant in der Ausbildung zukünftiger Lehrkräfte sind und sein können.

Effekte der Klassengröße auf Leistungen in Mathematik und Deutsch in der gymnasialen Sekundarstufe - Quantitative Analyse der VERA 8-Daten aus Baden-Württemberg

Maren Schubert[1,2], Karina Karst[2], Johannes Schult[3]
[1]Zentrum für internationale Bildungsvergleichsstudien (ZIB), Technische Universität München, Germany; [2]Universität Mannheim; [3]Institut für Bildungsanalysen Baden-Württemberg

Bei der Diskussion um Gründe für schlechteres Abschneiden von deutschen Schüler*innen wird nicht zuletzt die Klassengröße als Einflussfaktor thematisiert. Während überlastete Lehrkräfte auf Basis neuer Studien (Bach & Sievert, 2020) für niedrigere Klassenteiler plädieren (Baumann, 2022), bezog sich die Vorsitzende der Kultusministerkonferenz (KMK) Karin Prien im Jahr 2022 auf andere Studien (Hattie, 2005) und forderte zur Begegnung des akuten Lehrkräftemangels sogar eine Vergrößerung der Klassen (Gillmann, 2022). Diese Diskussionen sorgen auch bei Eltern, die Bildungsentscheidungen für ihre Kinder treffen, zu Unsicherheit.

Mit Blick auf international vergleichende Studien scheint evident, dass die Klassengröße kaum bis keinen Einfluss auf die Leistung der Schüler*innen hat (Hattie, 2009; Hoxby, 2000). Zuletzt fanden Bach und Sievert (2020) jedoch signifikant – wenn auch moderat – positive Effekte der Reduktion der Klassengröße auf Leistungen in saarländischen Grundschulen. Die Autoren gehen davon aus, dass Schüler* innen innerhalb einer Schule beispielsweise aufgrund ihrer Leistungen nicht zufällig in kleinere beziehungsweise größere Klassen eingeteilt werden. Dies sei bislang nicht berücksichtigt worden und könne bisherige Untersuchungen verfälscht haben (Bach & Sievert, 2020, S. 19).

Dieser Hypothese soll auch im vorliegenden Beitrag nachgegangen werden, indem die von Bach und Sievert (2020) vorgeschlagene und anhand VERA-3 Daten aus dem Saarland durchgeführten Auswertungsstrategie auf Gymnasialklassen in Baden-Württemberg angewendet wird. Hierfür wurde eine Sekundäranalyse der VERA 8-Daten aus den Jahren 2016 bis 2019 in Mathematik und Deutsch durchgeführt. Die Stichprobe umfasste für Mathematik N = 97.681 Schüler*innen aus 4.293 achten Klassen und für Deutsch N = 101.134 Schüler*innen aus 4.404 achten Klassen aller baden-württembergischen Gymnasien. Die durchschnittliche Klassengröße lag bei M = 25.35 (SD = 3.17).

Um die nicht zufällige Zuteilung von Schüler*innen in Klassen zu berücksichtigen, wurde wie folgt vorgegangen: Neben der faktischen Klassengröße wurde zunächst die durchschnittliche Klassengröße einer Schulkohorte berechnet. Zudem wurden die Leistungswerte (abhängige Variable) am schulübergreifenden Kohortenmittelwert z-standardisiert, um für jahrgangsspezifische Unterschiede der Messinstrumente zu kontrollieren. Zuletzt wurden die faktischen sowie durchschnittlichen Klassengrößen und Leistungswerte an den durchschnittlichen Werten der jeweiligen Schule – über alle Kohorten hinweg – group-mean-zentriert. Hierdurch wurde der Effekt der schul- und jahrgangsspezifischen Abweichung der faktischen (unabhängige Variable$_1$) sowie schul- und jahrgangsspezifischen durchschnittlichen Klassengröße (unabhängige Variable$_2$) von der schulspezifischen durchschnittlichen Klassengröße über alle Jahrgänge hinweg) auf die Abweichung der Leistung vom schulspezifischen Mittelwert (aV) geschätzt.

Die Analyse erfolgte in zwei Schritten:

1) Vorhersage der Mathematik- bzw. Deutschleistung auf Basis faktischer und durchschnittlicher Klassengröße mittels linearer Regressionen ohne zusätzliche Berücksichtigung der Mehr-Ebenen-Struktur und mit Schätzung von Erwartungswerten für unterschiedlich große Klassen. (in Anlehnung an Bach & Sievert, 2020)

2) Vorhersage der Mathematik- bzw. Deutschleistung mittels Mehr-Ebenen-Regression, wobei faktische und durchschnittliche Klassengröße als Merkmale der Klasse (Level 2) in die Analysen aufgenommen wurden.

Alle Analyseschritte zeigen keine signifikanten Effekte der Klassengröße auf die Ergebnisse in VERA 8. Auch die Verwendung der durchschnittlichen statt der faktischen Klassengröße wirkt sich nicht auf die Ergebnisse aus. Die Ergebnisse reihen sich somit in bestehende internationale Forschungsergebnisse ein. Eine Generalisierung der Befunde von Bach und Sievert (2020) auf baden-württembergische Gymnasien ist nicht möglich. Dies ist mit Blick auf höhere Klassenteiler an Gymnasien im Vergleich zu Grundschulen nachvollziehbar.

Die vorliegende Studie ergänzt und erweitert den Diskurs um die Klassengröße und liefert vorläufig Evidenz dafür, dass der Einfluss der Klassengröße in der öffentlichen Diskussion möglicherweise überschätzt wird. Zudem bietet sie Grundlage zur Diskussion, inwiefern im wissenschaftlichen Diskurs Analysestrategien weiter verfeinert werden können und darüber hinaus, wie Eltern und Politik durch transparente Kommunikation des aktuellen Forschungsstandes Unsicherheiten in diversen Entscheidungsprozessen genommen werden können.

C-06: Bildungshintergrund und Bildungserfolg

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S05 T00 B59
Chair der Sitzung: Emilija Meier-Faust

Unequal opportunities in high school – Interactions of socioeconomic background and personality in predicting school performance

Emilija Meier-Faust[1], Annelie Schulze[2], Yannick Martin[3], Annabell Daniel[4], Susanne Bergann[3]
[1]Deutsches Institut für Wirtschaftsforschung (DIW) Berlin; [2]Universität Potsdam; [3]Freie Universität Berlin; [4]Ludwig-Maximilians-Universität München

While socioeconomic background (Sirin, 2005) and personality (Poropat, 2009) have both been identified as relevant predictors of academic achievement, little is known about their possible interplay in their influence on students' performance. Academically relevant personality traits such as Conscientiousness might *compensate* for background disadvantages (*resource substitution hypothesis*) which has been found in some studies (Ayoub, Gosling, Potter, Shanahan, & Roberts, 2018; Damian, Su, Shanahan, Trautwein, & Roberts, 2015; Shanahan, Bauldry, Roberts, Macmillan, & Russo, 2014), however only with regard to adults' educational success. It is also possible that privileged students benefit more from advantageous traits (*Matthew effect hypothesis*). A recent German study (Lechner, Bender, Brandt, & Rammstedt, 2021), accordingly, reported positive interactions between parental socioeconomic status (SES) and ninth graders self-reported Conscientiousness. Personality and socioeconomic background, finally, might also have *independent* effects on academic achievement.

In order to amplify existing empirical knowledge, the present study introduced a process-based measure of socioeconomic background – cultural capital – in addition to SES as well as parent-reported personality as predictors of school performance in high school. Watermann and Baumert (2006) proposed that social inequalities in education might be better explained when incorporating different measures of socioeconomic background – *structural* background characteristics reflected by SES as well as *process-based* indicators that reflect cultural activities and cultural resources at home. We, accordingly, used both SES and cultural capital to investigate possible interaction effects of socioeconomic background and personality. With regard to the latter, different raters might provide unique information (Vazire, 2010) on individuals and particularly parents are considered good raters of their offspring's personality (Tackett, 2011). In a meta-analysis on primary education (Poropat, 2014), Conscientiousness and Openness showed stronger associations with achievement when adult-ratings were used compared to self-reports. Parental reports on personality, consequently, might be highly relevant predictors of school performance and, moreover, relevant with regard to possible interactions effects. Taken together, we investigated, whether 1) SES and cultural capital as well as 2) parent-rated Big Five predicted ninth graders' school performance and if 3) these effects were independent or interacting.

We employed structural equation modeling and used the Latent Moderated Structural Equations method (Klein & Moosbrugger, 2000) to investigate latent interactions between socioeconomic background measures and parent-rated Big Five when predicting school performance in large-scale data on German ninth graders (N = 2,770) from the National Educational Panel Study (NEPS, Blossfeld and von Maurice, 2011).

In line with previous empirical findings, SES was a significant predictor of high school students' school performance over and above fluid intelligence, gender, migration background, and school track. Cultural capital showed a marginal incremental effect over and above SES and mediated the effect of SES on school performance. Openness and Conscientiousness significantly predicted school performance over and above fluid intelligence, demographic control variables and socioeconomic background. We found evidence for significant positive interactions of those two traits with SES and cultural capital in support of the *Matthew effect hypothesis.* Students with a more privileged background benefited more strongly from their high Openness and high Conscientiousness. This evidence might be particularly relevant for planning school support programs and interventions targeting background disadvantages. Our study suggests that unequal opportunities in school might be targeted most efficiently when addressing both a lack of resources as well as academically beneficial personality characteristics.

Gleicher Abschluss, gleiche Note, aber unterschiedliche Arbeitsmarktchancen? – Experimentelle Evidenz zur schulformabhängigen Signalwirkung von Abschlussnoten bei der Bewerber*innenauswahl

Benjamin Schimke, Claudia Schuchart
Bergische Universität Wuppertal, Deutschland

Einleitung & Theorie

Bildung *zahlt* sich aus! Viele internationale Studien zeigen, dass mehr Bildung auch bessere Arbeitsmarktchancen bedeutet (z. B. Psacharopoulos und Patrinos 2004). Allerdings lassen einige Befunde annehmen, dass es bei gleichem Bildungsniveau auch darauf ankommt, in welche Bildung investiert wird (Birkelund et al. 2021). Dies ist besonders relevant für das deutsche Bildungssystem, welches sich durch eine frühe Gliederung in der Sekundarstufe I auszeichnet. Um dennoch Korrekturen einmal getroffener Schulartentscheidungen zu ermöglichen, gibt es vor allem in der Sekundarstufe II ein vielfältiges Angebot zum alternativen Abschlusserwerb, z.B. an beruflichen Schulen (Schuchart 2013). Allerdings scheint z.B. eine Hochschulreife, die auf einer beruflichen Schule erworben wurde, zu geringeren Arbeitsmarktrenditen zu führen als der gleichnamige Abschluss vom Gymnasium (Schuchart und Schimke 2019). Bislang ist unklar, ob diese Befunde auf Selbstselektion oder Selektion durch Arbeitgeber*innen zurückzuführen sind. Die zentrale Fragestellung dieses Beitrags richtet sich darauf, inwieweit der Weg zum Bildungsabschluss die Einladung zum Vorstellungsgespräch beeinflusst.

Die Basis des theoretischen Rahmens bildet die Signal- und Filtertheorie nach Spence (1973), Arrow (1973) und Stiglitz (1975). Demzufolge könnten verschiedene Wege zum gleichnamigen Bildungsabschluss mit unterschiedlicher Signalwirkung verbunden sein, wenn Absolvent*innen unterschiedlicher Schulformen über differenzielle produktivitätsrelevante Merkmale verfügen. So wurden z.B. schulformbezogene Leistungsdisparitäten zugunsten der Schüler*innen am Gymnasium nachgewiesen (Leucht et al. 2016). Eine bislang offene Frage ist, ob gleichnamige Abschlüsse und erbrachte Leistungen an unterschiedlichen Schulformen auch von Personaler*innen als Signal unterschiedlicher Produktivität bewertet werden.

Methodik

Die Studie basiert auf einem faktoriellen Survey (Auspurg und Hinz 2015). Das Experiment wurde im Januar 2022 mit Personaler*innen (N=717) durchgeführt, die Bewerber*innen in mindestens einem von 15 Berufen rekrutieren. Die Berufe wurden ausgewählt um, 1) die berufliche Geschlechtersegregation, 2) unterschiedliche Anforderungsniveaus und 3) quantitativ bedeutsame Berufe am deutschen Arbeitsmarkt abzubilden. Zu Beginn des Experiments wurden jedem Probanden eine fiktive Stellenausschreibung aus jenem Beruf präsentiert, in dem sie selbst Einstellungsexpertise besitzen. Im Anschluss wurden die Probanden gebeten nacheinander insgesamt acht tabellarische Lebensläufe fiktiver Bewerber*innen zu bewerten. Hierzu stand jeweils eine 11-stufige Skala zur Verfügung mit der die Wahrscheinlichkeit einer Einladung zu einem Vorstellungsgespräch eingeschätzt werden sollte. Die experimentell variierten Faktoren umfassten u.a. das Geschlecht, die Abschlussnoten sowie ein dreistufiger Faktor mit Schulformen des höchsten allgemeinbildenden Schulabschlusses (Gymnasium, Gesamtschule und berufliche Schule). Das experimentelle Design umfasst insgesamt 10 Faktoren, ist strukturell identisch für alle 15 Berufe und basiert auf 192 verschiedenen Lebensläufen, die experimentell aus dem Vignettenuniversum ($2^6 3^4$) (d-Effizienz: 92,8%) ausgewählt und anschließend experimentell 24 Vignettensets zugewiesen wurden, um die Konfundierung aller Haupteffekte und aller 2-fach Interaktionen zu verhindern. Die Analyse der bewerteten Vignetten (N*V=5.443) erfolgt mittels hierarchisch-linearer Mehrebenenregression mit random intercepts.

Ergebnisse

Die Analysen zeigen, dass Absolvent*innen mit weniger guten Abschlüssen (1,3 vs. 3,0) eine statisch signifikant geringere (-2,6 Prozentpunkte) Einladungswahrscheinlichkeit aufweisen. Gleichnamige Abschlüsse, die auf unterschiedlichen Schulformen erworben wurden, sind nicht direkt mit Vor- oder Nachteilen in der frühen Phase von Bewerbungsprozessen verbunden. Die Signalwirkung von Abschlussnoten variiert jedoch schulformabhängig. Im Vergleich zu einer sehr guten Abschlussnote (1,3) vom Gymnasium, werden Absolventen von Gesamtschulen (bei gleicher Note) mit signifikant geringerer Wahrscheinlichkeit zu Vorstellungsgesprächen eingeladen. Die Unterschiede sind insgesamt gering (-1,8 Prozentpunkte), jedoch kausal zu interpretieren. Für weniger gute Abschlussnoten (3,0) lassen sich die Resultate nicht reproduzieren.

Die Ergebnisse sind bemerkenswert, da die fiktiven Stellenausschreibungen an Bewerber*innen mit Berufserfahrung gerichtet und diese konstant in allen Vignetten mit zwei Jahren angegeben ist. Darüber hinaus weisen alle Vignetten einen entsprechend notwendigen beruflichen Abschluss / ein Hochschulstudium und deren Abschlussnoten auf. Neben diesen Arbeitsmarktsignalen scheint der Schulform des allgemeinbildenden Abschlusses eine weitere Rolle zuzukommen. Insbesondere für die Bildungsberatung von Schüler*innen sowie deren Eltern, Auszubildenden und Studierenden besitzen die Erkenntnisse eine erhebliche Relevanz.

Bildungserfolg durch Separation? Befunde zum Programm „Produktives Lernen in Schule und Betrieb“

Raphaela Porsch, Melanie Baumgarten, Jennifer Quast, Robert W. Jahn
Otto-von-Guericke Universität Magdeburg, Deutschland

Fachdidaktiker*innen und Schulpädagog*innen bemühen sich seit Jahren um die Entwicklung und Erforschung von Konzepten, um den Schulerfolg aller Schüler*innen mit unterschiedlichen Lernvoraussetzungen zu ermöglichen. Neben didaktischen Konzepten zur individuellen Förderung innerhalb des Unterrichts, bspw. im Kontext inklusiven Lernens (Korff, 2012), liegen Programme vor, die einzelne Schüler*innen separieren. Als „Ausnahme von der Regelbeschulung“ bezeichnet (MK, 2011), richtet sich bspw. das Programm „Produktives Lernen (in Schule und Betrieb)“ (PL) an Schüler*innen, „bei denen zu erwarten ist, dass sie im Regelsystem nicht mindestens einen Hauptschulabschluss erreichen werden“ (ebd.). Es wird aktuell in Berlin, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt und Schleswig-Holstein angeboten. Schüler*innen lernen innerhalb einer Woche in einem praktisch ausgerichteten Unterricht in der Schule und in einem Betrieb. Zentrale Merkmale sind die Individualisierung von Lernprozessen, der Erwerb von fachlichen und überfachlichen Kompetenzen in der Berufspraxis und im Beruf sowie die Verbindung von (funktionaler) Allgemeinbildung und Berufsorientierung (vgl. Borkenhagen & Mirow (2002, S. 444). Obwohl es bereits in den 1980er Jahren entwickelt wurde (Bliss, 2008), liegen kaum belastbare empirische Befunde zur Wirksamkeit des Programms vor. Hauptziel des PL-Programms ist es, zum ersten anerkannten Schulabschluss zu führen. Um dieses Ziel zu erreichen, soll in den Klassenstufen 8 und 9 eine Verringerung des Schulschwänzens und der Anzahl an Klassenwiederholungen erreicht werden. Die schulischen Leistungen sollen verbessert und ausbildungsrelevante Kompetenzen erworben werden. Vor diesem Hintergrund sollen folgende Fragen beantwortet werden: (1) Lässt sich im Vergleich zu Regelschüler*innen eine Verringerung von Schulabsentismus und Klassenwiederholungen feststellen? (2) Verbessern sich die Schulleistungen der PL-Schüler*innen im Verlauf der Teilnahme des Programms? (3) Wie hoch ist die Erfolgsquote der PL-Schüler*innen, d.h. derjenigen, die einen ersten Schulabschluss erreichen? Obwohl es kein direktes Ziel des PL-Programms ist, erhalten einige PL-Schüler*innen am Praxislernort einen Ausbildungsvertrag. Vor diesem Hintergrund stellt sich die Frage (4) Wie hoch ist die Anzahl von Neuntklässler*innen – Regelschüler*innen vs. PL-Schüler*innen, die bereits einen Ausbildungsvertrag besitzen oder sich bereits beworben haben? Zur Beantwortung liegen Daten einer Befragung in Sachsen-Anhalt von 451 PL-Schüler*innen und 1.729 Regelschüler*innen der 8. und 9. Klassen aus dem Schuljahr 2021/22 vor. Für die Auswertung bzw. die Gruppenvergleiche wurden inferenzstatistische Analysen durchgeführt (Chi-Quadrat Tests o. t-Tests).

Die Ergebnisse zeigen u.a.: PL-Schüler*innen ($M = 1,05$; $SD = 1,21$; $N = 451$) gaben an, in ihren ersten sieben Schuljahren signifikant häufiger ($t = 6,683$, $p < .001$) die Schule geschwänzt zu haben als Mitschüler*innen der regulären 8. und 9. Klassen ($M = 0,65$; $SD = 0,83$; $N = 1.729$). Hingegen unterscheiden sich PL-Schüler*innen ($M = 0,81$; $SD = 1,79$; $N = 451$) und Regelschüler*innen ($M = 0,67$; $SD = 1,48$; $N = 1.729$) im aktuellen Schuljahr nicht hinsichtlich des Schwänzens ($t = 1,591$, $p = .112$). PL-Schüler*innen sind nach eigener Aussage nicht nur häufiger von Klassenwiederholungen betroffen, sie mussten auch eine größere Anzahl von Klassenstufen wiederholen ($Z = 9.735$, $p < .001$). Die Teilnahme am PL-Programm bringt dann eine deutliche Veränderung: Ist der Anteil in der 8. Klasse an Wiederholer*innen im PL noch größer als im Regelschulsystem, sind in Klasse 9 signifikant weniger PL-Schüler*innen von Klassenwiederholungen betroffen. Für die Frage der

erfolgreichen Teilnahme wurden für die letzten sechs Schuljahre Erfolgsquoten zwischen 76,65% und 85,19% ermittelt. Die Dropoutquoten liegen zwischen 9,37% und 18,51% und weisen einen Gesamtmittelwert von 12,81% auf. Die durchschnittliche Dropoutquote an weiterführenden Schulen in Sachsen-Anhalt liegt bei rund 10% (Porsch, Jahn & Baumgarten, 2021). Wird berücksichtigt, dass ausschließlich von Schulabbruch gefährdete Schüler*innen am PL-Programm teilnehmen, wird deutlich, dass das Programm erfolgreich Schüler*innen zum Schulabschluss führt. Neben der Vorstellung dieser und weiterer Ergebnisse zur Wirksamkeit des Programms soll abschließend die Frage diskutiert werden, ob eine Separation von Schüler*innen, die über Risikomerkmale verfügen, pädagogisch sinnvoll ist bzw. was dafür oder dagegen spricht.

Schwierige Ausgangslagen für Bildungserfolg und individuelle Entwicklung: Inwieweit kumulieren soziale und individuelle Risikofaktoren am Ende der Grundschule?

Elena-Christin Haas[1], Marko Neumann[1], Jan Scharf[1], Michael Becker[1,2], Kai Maaz[1], Jürgen Baumert[3], Olaf Köller[4]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main/Berlin; [2]Institut für Schulentwicklungsforschung, TU Dortmund, Dortmund; [3]Max-Planck-Institut für Bildungsforschung, Berlin; [4]IPN - Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel

Zum Ende der Primarschulzeit sollen für alle Kinder die Grundlagen für den Besuch der weiterführenden Schule gelegt sein. Bildungsauftrag der Grundschule ist es, allen Kindern, unabhängig ihrer familiären und sozialen Herkunftsbedingungen und Voraussetzungen, sowohl in ihren schulischen Leistungen als auch in ihrer Persönlichkeitsentwicklung zu stärken und zu fördern (vgl. Martschinke, 2019). Wie zahlreiche Untersuchungen belegen, wird dieser Anspruch jedoch nur begrenzt eingelöst. Frühe Bildungsungleichheiten – so auch am Ende der Grundschulzeit – können dabei weitreichende Folgen im Bildungs- und Lebensverlauf haben (vgl. u.a. Blossfeld et al., 2019; Maaz et al., 2006), stellt Bildung doch eine grundlegende Voraussetzung für gesellschaftliche Teilhabe, schulischen und beruflichen Erfolg aber auch gesundheitliches Wohlergehen dar (vgl. Hadjar et al., 2008). Nationale und internationale Studien aus dem Kontext der Bildungsungleichheitsforschung zeigen robuste herkunftsbedingte Disparitäten im Bildungserfolg zu Ungunsten von Schüler*innen aus sozial weniger privilegierten Familien bzw. mit schwierigen familiären Ausgangslagen auf. Diese sind u.a. bereits bei frühen Bildungsübergängen (vgl. Doblinger, 2020; Dumont et al., 2019), schulischen Kompetenzständen in Grundschule und Sekundarstufe (vgl. Neumann et al., 2014) und damit zusammenhängend auch für den nachschulischen Bildungsbereich bis zum Einstieg ins Berufsleben (vgl. Hupka-Brunner et al., 2015) feststellbar. Familiäre Herkunftsdisparitäten finden sich auch hinsichtlich psychosozialer bzw. sozioemotionaler Merkmale, etwa der lernbezogenen Motivation oder Lesefreude der Schüler*innen (vgl. Hertel et al., 2010; vgl. Hillmayr et al., 2021; Petermann & Weber, 2016).

Trotz einer umfassenden Forschungsbasis zu den Faktoren (nicht) gelingender Bildungs- und Entwicklungsverläufe bleibt in der wissenschaftlichen und öffentlichen Diskussion oftmals eher vage, was unter „schwierigen Ausgangslagen“ oder „Risikolagen“ im Sinne ungünstiger Startbedingungen verstanden werden kann. Dies gilt insbesondere mit Blick auf die Kumulation und Kombinationen von Einzelfaktoren bzw. unterschiedlichen Facetten schwieriger Ausgangslagen. Eine facettenreiche Betrachtung bildungsbezogener schwieriger Ausgangslagen ist jedoch relevant, da davon auszugehen ist, dass sie kumulierend und in Abhängigkeit ihrer Kombination mögliche Konsequenzen

in der Entwicklung nach sich ziehen (vgl. Rutter, 1979; Scheithauer & Petermann, 1999; Tetzner, 2014). Für den Bildungsbereich liegen insbesondere in der Betrachtung verschiedener Kombinationen bzw. Konfigurationen bildungsbezogener Risikolagen bislang nur wenige Untersuchungen vor. Diese sind jedoch wichtig, um spezifische Bedürfnisse von Schüler*innen erkennen und passende Unterstützungsansätze entwickeln zu können (vgl. Fortin et al., 2006; Lorenz et al., 2020).

An dieser Stelle setzt der vorliegende Beitrag an. Im Zentrum steht dabei die Frage, inwieweit sich unterschiedliche Facetten potenzieller bildungsbezogener schwieriger Ausgangslagen kumulieren und in bestimmten Konfigurationen auftreten. Abgeleitet von entwicklungspsychologischen Theorien zur Wirkweise von Risikofaktoren sowie bildungssoziologischen Wirkmodellen von familiären Lebensstrukturen auf den Bildungserwerb (vgl. Baumert et al., 2003; Hasselhorn et al., 2017; McElvany et al., 2009; Scheithauer & Petermann, 1999), werden potenzielle bildungsbezogene Risikofaktoren definiert und differenziert in Form von familiären Strukturmerkmalen (z.B. elterliche Bildung, Erwerbstätigkeit) und Prozessmerkmalen (z.B. elterliche Unterstützungsmöglichkeiten) sowie individuellen schulbezogenen und sozioemotionalen Schüler*innenmerkmalen (Noten, Klassenwiederholung als kritisches schulbiographisches Ereignis, schulische Ängste und Überforderung) analysiert. Die Datengrundlage bildet eine repräsentative Schüler*innenstichprobe (*N* = 3935) am Ende der sechsten Jahrgangsstufe im Land Berlin (BERLIN-Studie, vgl. Maaz et al., 2013). Mithilfe latenter Klassenanalysen werden verschiedene Konfigurationen von schwierigen Ausgangslagen am Ende der Grundschule ermittelt, beschrieben und in ihrer Vorkommenshäufigkeit untersucht.

Die latenten Klassenanalysen ergaben drei Gruppen von Schüler*innen mit unterschiedlich hohen Wahrscheinlichkeiten (multipler) schwieriger Ausgangslagen vor dem Übergang in das gegliederte Schulsystem: 27 Prozent der Schüler*innen ließen sich einer Gruppe zuordnen, die durchgängig in allen betrachteten Bereichen hohe Risikowahrscheinlichkeiten aufwies. Etwa 19 Prozent der Schüler*innen bildeten eine Gruppe, die trotz hoher familiärer Risikowahrscheinlichkeiten nur geringe individuelle Risikowahrscheinlichkeiten aufwies. 54 Prozent fielen in die Gruppe mit durchgängig geringen Risikowahrscheinlichkeiten. Die resultierende Drei-Klassen-Lösung erwies sich auch in einer Modelltestung mittels Split-Half-Designs als robust. Forschungsbezogene und praktische Implikationen des gewählten personenzentrierten Ansatzes und der Ergebnisse werden diskutiert.

C-07: Hochschulbildung

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S06 S04 B06
Chair der Sitzung: Julia Steinkühler

Warum ist ein Studienabbruch durch die soziale Herkunft geprägt?

Julia Steinkühler, Sören Isleib
DZHW, Germany

Angesichts stagnierend hoher Abbruchquoten von rund 28 % im Bachelorstudium (Heublein et al. 2022) sind die Erklärungen möglicher Ursachen und anschließender Konsequenzen des Studienabbruchs im öffentlichen und wissenschaftlichen Diskurs ein relevantes Thema der Hochschulforschung. Bisher wurden hohe Abbruchquoten hauptsächlich als volkswirtschaftlicher Verlust (potentielle Fachkräfte, Innovationspotentiale) oder ineffizienter Einsatz finanzieller Mitteln auf Hochschulebene gewertet (Sarcletti & Müller 2011). Gleichzeitig orientiert sich die Mittelvergabe an Hochschulen zum Teil an ebenjenen Performanzkriterien (Thomas & Hovdhaugen 2014). Der wissenschaftliche Blick auf den Studienabbruch folgt eher der Bestimmung individueller Abbruchrisiken, wobei zunehmend ungleichheitstheoretische Fragestellungen im Mittelpunkt stehen. Angesichts steigender Studierendenzahlen und einer allgemein stärkeren Bildungsbeteiligung ehemals bildungsferner Gruppen stellt sich die Frage, inwieweit Öffnungsprozesse in frühen Bildungsstufen (u.a. Übergang in den Sekundarbereich, Erwerb der Hochschulzugangsberechtigung) mit unterschiedlichen Erfolgschancen in späten Bildungsstufen (Studium) verbunden sind (Schindler 2015).

Im Kontext der Ungleichheitsforschung erweist sich die soziale Herkunft für viele Bildungsentscheidungen und –prozesse als distinktives Merkmal. Der vorgeschlagene Beitrag untersucht, ob und warum die soziale Herkunft einen Einfluss auf den Studienabbruch hat. Die Studienphase stellt eine späte Bildungsstufe im individuellen Bildungsverlauf dar, welcher bereits verschiedene, herkunftsspezifisch geprägte Bildungsübergänge vorausgehen (Boudon 1974; Dumont et al. 2014). Da Studierende aus einem nicht-akademischen Elternhaus in besonderem Maße berufspraktische Wege zur Hochschulzugangsberechtigung frequentieren und im Vergleich schlechtere Schulleistungen aufweisen, beide Faktoren jedoch mit erhöhten Abbruchwahrscheinlichkeiten einhergehen (Müller & Schneider 2013), wird davon ausgegangen, dass diese Merkmale einen Studienabbruch begünstigen. Daneben folgt der Beitrag einem prozesshaften Verständnis des Studienabbruchs und fragt nach dem Erklärungspotential etwaiger herkunftsspezifischer Differenzen beim Übergang in die Hochschule (Hochschul-/Studienfachwahlmotive) und im Studienprozess (soziale/akademische Integration bzw. Passung zwischen individuellem und akademischem Wertesystem) (Tinto 1975; Bourdieu & Passeron 1977). Herkunftsspezifische Unterschiede können darüber hinaus in studienexternen Faktoren, wie Finanzierungsicherheit und Erwerbstätigkeit, begründet sein (Brandstätter & Farthofer 2003; Choi 2017).

Der vorliegende Beitrag untersucht die Frage, inwieweit im Abbruch eines Bachelorstudiums herkunftsspezifische Disparitäten bestehen und wodurch sich diese Unterschiede erklären lassen. Dabei werden die oben beschriebenen bildungsbiografischen und hochschulischen sowie studienprozess- und studienexternen Herkunftseffekte berücksichtigt. Grundlage der Analysen bildet

die Exmatrikuliertenstudie des DZHW (Heublein et al. 2017), in der ca. 9.100 Studierende drei bis sechs Monate nach ihrer Exmatrikulation befragt wurden. Zunächst wird mittels logistischer Regressionsmodelle der Einfluss der sozialen Herkunft auf den Studienabbruchs geschätzt (average marginal effects). Über eine anschließende nicht-lineare Dekompositionsanalyse nach Karlson et al. (2012) wird ermittelt, in welchem Ausmaß die Herkunftsunterschiede im Abbruchverhalten auf die theoretisch extrahierten Merkmale zurückzuführen sind.

Es zeigt sich, dass Studierende aus einem nicht-akademischen Elternhaus eine zunächst signifikant erhöhte Abbruchwahrscheinlichkeit aufweisen, bei simultaner Kontrolle aller potenziellen Einflussfaktoren jedoch keine signifikanten Herkunftsdifferenzen mehr bestehen. Dieser Befund wirft die Frage auf, welche konkreten Erklärungsanteile die einzelnen Faktoren an den bestehenden herkunftsspezifischen Disparitäten hinsichtlich des Studienabbruchs aufweisen. Mit der Dekompositionsanalyse kann gezeigt werden, dass insgesamt 76 % des herkunftsbedingten Unterschieds durch die untersuchten Faktoren erklärt werden können. Die höhere Abbruchwahrscheinlichkeit statusniedriger Gruppen lässt sich vor allem auf die Wahl der Hochschulart, nicht-gymnasiale Zugangswege an die Hochschule, ein geringeres Maß an sozialer und akademischer Integration sowie größere Finanzierungsschwierigkeiten zurückführen. Mit den Ergebnissen können somit Bildungsungleichheiten und deren Mechanismen an einer späten Gelenkstelle des Bildungssystems nachgewiesen werden.

Training von interprofessionellen Kooperationsfertigkeiten – Ein Vergleich der Effekte von externalen Kooperationsskripts und Gruppenreflektion

Julia Eberle, Sebastian Strauß, Isis Tunnigkeit, Leonie vom Bovert, Marcel Schmittchen, Arlind Avdullahu, Nikol Rummel
Ruhr-Universität Bochum, Germany

Interprofessionelle Kooperation prägt zunehmend die Arbeitswelt in vielen Bereichen. Beispielsweise erfordern Großbauprojekte die Zusammenarbeit von Bauingenieuren, Architekten und Brandschutzexperten. Ein Scheitern der interprofessionellen Kooperation ist hier mit schwerwiegenden Konsequenzen verbunden. Das Gelingen interprofessioneller Kooperation ist jedoch komplex und erfordert den Austausch ungeteilter Informationen, das Herstellen von Verbindungen zwischen diesen Informationen, das Aushandeln oft widersprüchlicher Lösungsansätze sowie das Treffen komplexer Entscheidungen (Brodbeck et al., 2007).

Zur Vermittlung von Kooperationsfertigkeiten hat sich in zahlreichen Studien die Verwendung von externalen Kooperationsskripts bewährt (Radkowitsch, Vogel, & Fischer, 2020). Bei diesem Ansatz werden im Verlauf eines Kooperationsprozesses günstige Handlungsanweisungen geprompted, die von den Lernenden internalisiert werden sollen. Allerdings kann der stark direktive Charakter dieser Instruktionsmaßnahme ein Nachteil sein. Ein weniger direktiver Ansatz (Rummel, 2018) ist der Einsatz von Group-Awareness-Tools (GAT), die die Verteilung von Merkmalen in einer Gruppe, wie etwa der Zufriedenheit mit einem Aushandlungsprozess, sichtbar machen. GATs haben sich für die computergestützte Arbeit bereits bewährt (Schmidt & Randall, 2016) und ihr Potenzial wird zunehmend auch für das kooperative Lernen untersucht (Bodemer, Janssen, & Schnaubert, 2018). Hierbei bietet sich insbesondere eine Kombination mit einer gemeinsamen Reflektionsphase an.

Erkenntnisse über die Wirksamkeit dieses Ansatzes für die Verbesserung von Kooperationsfertigkeiten sowie die daraus resultierende kognitive Belastung fehlen jedoch bislang.

Forschungsfragen:

1. Inwiefern wirken sich keine Unterstützung, ein Kooperationsskript oder ein GAT mit Reflektion in einer kooperativen Problemlösephase auf den Erwerb von Kooperationswissen aus?

2. Inwiefern unterscheidet sich die wahrgenommene kognitive Belastung in einer kooperativen Problemlösephase ohne Unterstützung, mit Kooperationsskript oder mit GAT mit Reflektion?

In einer Online-Laborstudie mit 150 Studierenden beschäftigten sich 50 Dreiergruppen mit jeweils zwei Problemszenarien, die eine vereinfachte interprofessionelle Kooperationssituation aus dem Bauingenieursbereich simulierten. Die Dreiergruppen wurden zufällig einer von drei Bedingungen zugeordnet (Kontrollbedingung, Kooperationsskript, GAT). Jede Person erhielt eine Rollenbeschreibung, die ungeteilte Informationen, individuelle Anforderungen und Lösungsansätze enthielt, die für die Lösung der Problemszenarien in der Dreiergruppe als Diskussionsgrundlage dienten.

In allen Bedingungen füllten die Versuchspersonen zunächst einen Vorfragebogen mit Wissenstest aus. Darauf folgte die erste Problemlösephase (Lernphase), in der a) die Dreiergruppen in der Kontrollbedingung keine Unterstützung erhielten, b) die Gruppen in der einen Experimentalbedingung ein Kooperationsskript erhielten, das die Problemlösephase vorstrukturierte und Prompts für günstige Strategien und deren Anwendung umfasste und c) die Gruppen in der zweiten Experimentalbedingung unangeleitet das Problem bearbeiteten und im Anschluss eine Gruppenreflektion anhand eines GATs durchführten, wobei das GAT die Einschätzung der Gruppenmitglieder zu zentralen Aspekten der Kooperationsqualität in der vorangegangenen Problemlösephase visualisierte. Es folgten für alle Bedingungen ein Zwischentest mit erneutem Wissenstest und einer kognitiven Belastungs-Einschätzung aus 2 Items. In diesem Paper unberücksichtigt bleiben die folgende zweite unangeleitete Problemlösephase (Testphase) sowie ein Abschlussfragebogen. Grundlage für die Erstellung von Kooperationsskript und GAT bildeten die neun Dimensionen der Kooperationsqualität von Meier, Spada, und Rummel (2007), z.B. gegenseitige Verständnissicherung, Gesprächssteuerung und Information Pooling. Die Auswertung erfolgte mittels einer ANOVA mit Messwiederholung und geplanten Kontrasten.

Bezüglich des Kooperationswissens zeigte sich ein signifikanter Interaktionseffekt zwischen Messzeitpunkt und Versuchsbedingung ($F(2, 147)=3{,}578$; $p<.05$). Eine erste Kontrastanalyse zeigte einen signifikanten Unterschied ($t(147)=3{,}231$; $p<.01$) zwischen der Bedingung ohne Unterstützung ($M=4{,}91$; $SD=2{,}35$) und den beiden Experimentalbedingungen mit Unterstützung. Eine zweite Kontrastanalyse zeigte zudem einen Unterschied ($t(147)=2{,}422$; $p<.05$) zwischen den beiden Experimentalbedingungen, wobei die Gruppe mit Kooperationsskript ($M=5{,}76$; $SD=2{,}55$) einen geringeren Wissenszuwachs aufweist als die Gruppe mit GAT und Reflektion ($M=6{,}94$; $SD=2{,}57$). In Bezug auf die kognitive Belastung während der Lernphase zeigten sich keine signifikanten Unterschiede zwischen den Gruppen.

Diese Ergebnisse zeigen das Potenzial von GAT mit Reflektion als einem wenig direktiven Ansatz für den Erwerb von Kooperationswissen, gegenüber dem Kooperationsskript-Ansatz. Die Überlegenheit bezüglich Wissenserwerb und fehlende Unterschiede hinsichtlich der kognitiven Belastung sind vielversprechend. Weitere Analysen von Prozessdaten sind nötig, um das Potenzial tiefergehend zu untersuchen.

Problemorientiertes Lehren und Lernen in der Hochschule: Validierung eines Fragebogens zur Erfassung der Unterstützung kooperativer Lernprozesse

Irmela Blüthmann[1], Anabel Bach[1], Ines Wulff[2], Harm Peters[2], Rainer Watermann[1], Felicitas Thiel[1]
[1]Freie Universität Berlin, Germany; [2]Charité - Universitätsmedizin Berlin

Problemorientiertes Lernen (POL) ist im Medizinstudium weltweit verbreitet. Über die Bearbeitung authentischer Fälle aus dem ärztlichen Berufsalltag lernen Studierende Fachinhalte, Problemlösen und Kooperation. POL findet in Kleingruppen weitgehend selbstgesteuert statt, entlang vordefinierter Schritte und unter der Anleitung von Tutor:innen. Dabei können zwei Quellen der Unterstützung kooperativer Lernprozesse unterschieden werden: Zunächst und in erster Linie soll die Unterstützung durch die Gruppe selbst erfolgen. Aufgabe der Tutor:innen ist es, den Lernprozess zu monitoren und nur bei Bedarf adaptiv zu unterstützen.

Die Wirksamkeit von POL belegen Metaanalysen (Dochy et al., 2003). Der Gruppenprozess selbst wurde vielfach qualitativ anhand von Fallbeispielen mit Fokus auf die Anregung der Elaboration durch *hints* und *prompts* (Hmelo-Silver 2004) untersucht. Es existieren jedoch kaum standardisierte Instrumente, die gleichzeitig beide Quellen der Unterstützung (Gruppe/Tutor:in) erfassen und Unterstützungsleistungen umfassend abbilden. Ausgehend von dem Modell der drei Basisdimensionen der Lehrqualität aus der Unterrichtsforschung (Praetorius et al. 2018) und dessen Adaption für die Hochschullehre (Thiel et al. 2012) haben wir ein neues Instrument entwickelt, das die kooperativen Lernprozesse in der Gruppe sowie deren Anregung durch Tutor:innen (hier: Lehrende) in diesen drei Dimensionen (Unterstützung der Wissenskonstruktion, der sozialen Interaktion und Motivierung) erfasst. Unter Rückgriff auf das Modell selbstgesteuerten Lernens von Schmitz & Schmidt (2007) bilden wir zusätzlich den Prozesscharakter bei POL ab und unterscheiden eine präaktionale, aktionale und postaktionale Phase. In der aktionalen Phase, in der die eigentliche Lernhandlung stattfindet, wurde mit 19 Items auf 4-stufigen Antwortskalen das Gelingen der kooperativen Prozesse in sechs Subdimensionen erfasst: Elaboration, Organisation, Metakognition, Lernatmosphäre, aufgabenbezogene Kommunikation und Partizipation (z.B. „*Unserer Gruppe ist es sehr gut gelungen, alle Gruppenmitglieder an der Diskussion zu beteiligen.*“). Mit 19 identischen Items mit anderem Intro („*Der/die Tutor:in hat uns dazu aufgefordert, …*“) wurde die Unterstützung durch die Tutor:innen in denselben Dimensionen erfasst.

Basierend auf den Daten eines Pretests an der Charité – Universitätsmedizin Berlin im Wintersemester 2020/21 (*N*=421 Studierende in 151 POL-Gruppen, 1.- 5. Fachsemester, Rücklauf: 25%) wurden verschiedene Maße der Konstrukt- und Kriteriumsvalidität berechnet. Mittels zwei konfirmatorischer Faktorenanalysen wurden die Faktorstruktur für beide Quellen der Unterstützung überprüft und die Interraterreliabilität (ICC) in allen Subdimensionen berechnet. Weiterhin wurden

Zusammenhänge auf Gruppenebene zwischen den Ratings und den Lernergebnissen (selbsteingeschätzter Kompetenzzuwachs und Zufriedenheit mit der Zusammenarbeit) berechnet.

Mit Blick auf die Konstruktvalidität zeigen die Ergebnisse, dass sich die 6 Subdimensionen in beiden CFA-Modellen gut abbilden lassen; allerdings waren Elaboration und Metakognition sehr hoch korreliert. Die Reliabilitäten sind zufriedenstellend bis sehr gut (.55<α<.92). Die Intraklassenkorrelationen sind für die Einschätzungen der Anregung durch die Tutor: innen (.19<ICC(1)<.36) etwas höher als für das Gelingen der Gruppenprozesse (.11<ICC(1)<.26). Auf Gruppenebene zeigen sich signifikante positive Zusammenhänge mit dem Kompetenzzuwachs und der Zufriedenheit mit der Zusammenarbeit, auch nach Kontrolle des Fachsemesters. Zudem korreliert die Anregung durch die Tutor:innen in allen Subdimensionen signifikant positiv mit dem Gelingen der Gruppenprozesse in der jeweiligen Subdimension. Scatterplots zeigen, dass es nur wenige Gruppen gibt, in denen die Prozesse auch ohne entsprechende Unterstützung gut gelingen. Die Ergebnisse sprechen insgesamt für die Validität des Instruments und verdeutlichen die Relevanz aller Subdimensionen bei POL. Weiterhin belegen sie die Bedeutung der POL-Tutor:innen. Dies ist aus zwei Gründen besonders erfreulich: 1. Da bei POL die Unterstützung auf sehr subtile Art und Weise erfolgt, stellt deren Operationalisierung eine besondere Herausforderung dar. 2. Da der Pretest des für die Präsenzlehre entwickelten Instruments im digitalen Setting erfolgen musste.

Eine Limitation stellt die geringe Clustergröße dar, die dazu führt, dass die Reliabilität der auf Gruppenebene aggregierten Werte (ICC(2)<.70) nicht zufriedenstellend ist. Dennoch zeigen sich erwartungskonforme Zusammenhänge, die für die kriteriale Validität sprechen. Eine Validierung in der Präsenzlehre ist geplant. Bei der Überarbeitung des Fragebogens ist die Operationalisierung der metakognitiven Prozesse noch zu verbessern.

Aktive Pausen an der Hochschule? – Effekte des „Pausenexpresses“ in Hochschulseminaren auf Aufmerksamkeit und Befindlichkeit

Christian Müller[1], Sven Lindberg[2], Jil Wolf[3], Marianne Schüpbach[3]
[1]Institut für angewandte Familien-, Kindheits- und Jugendforschung an der Universität Potsdam; [2]Universität Paderborn, Deutschland; [3]FU Berlin, Deutschland

Theoretischer Hintergrund

Mehr und mehr Hochschulen bieten Bewegungspausen in Lehrveranstaltungen zur aktiven Regeneration kognitiver Leistungsfähigkeit und zur Gesundheitsförderung ihrer Studierenden an (für eine Übersicht siehe z. B. Allgemeiner deutscher Hochschulverband, 2022). Die Effekte von Bewegungspausen, insbesondere von Bewegungspausen im Hochschulunterricht, auf kognitive Leistungen und motivationale Lernbedingungen sind bisher jedoch unzureichend untersucht (Infantes-Paniagua et al., 2021). Bisherige Befunde zu Effekten von körperlicher Aktivierung auf kognitive Fähigkeiten verweisen auf insgesamt positive Effekte auf Aufmerksamkeit, exekutive Funktionen und akademische Leistungen (Haverkamp et al., 2020), aber auch auf meta-analytische Befunde, die keine Effekte oder heterogene Befunde berichten (Amatriain-Fernández, Ezquerro García-Noblejas & Budde, 2021; Ishihara, Drollette, Ludyga, Hillman & Kamijo, 2021). Als Begründung hierfür werden methodische Schwächen, eine hohe Heterogenität bisheriger Studien angegeben und dass bisher

kaum Studien zu Bewegungspausen im Unterricht, bzw. Hochschulunterricht vorliegen (Müller et al., 2021; Singh et al., 2019; Wassenaar et al., 2020).

Fragestellung

Ziel der vorliegenden Interventionsstudie war es daher zu untersuchen, ob sich eine Bewegungspause im Hochschulunterricht positiver auf die Lernbedingungen Aufmerksamkeit und Befindlichkeit auswirken als konventionelle Pausen ohne körperliche Aktivierung.

Methode

Insgesamt nahmen 102 Studierende aus 4 inhaltsgleichen wöchentlichen universitären Seminaren beim gleichen Dozenten an der Studie teil (76,4 % = weiblich, *M*age = 24,29 Jahre, *SD* = 3.38).

Je 2 Seminare zählten zur Interventions- bzw. Kontrollgruppe. Als Bewegungspause wurde der „Pausenexpress“ einer Universität in einer deutschen Großstadt untersucht, bei dem geschulte Trainer:innen in die Lehrveranstaltungen kommen und mit den Studierenden eine freiwillige Bewegungspause durchführen. Die Bewegungspause dauerte jeweils 10 Minuten und startete stets nach 45 Minuten Seminarzeit. Die Bewegungspause hatte eine standardisierte Grundstruktur (Warm-Up, Übungsphase, Cool-Down) und orientierte sich an meta-analytischen Befunden hinsichtlich Dauer und Inhalten von Bewegungspausen (z. B. Mobilisations-, Dehnungs- und Entspannungsübungen; Singh et al., 2019). Zur Implementierung der Intervention wurden mehrere Wochen vor Beginn der Interventionsphase die Trainerin vorgestellt und die Bewegungspause im Seminar probeweise durchgeführt. In der Kontrollgruppe wurde eine konventionelle Pause oder keine Pause durchgeführt.

Als Pre- und Posttests wurden selektive Aufmerksamkeit (d2-R; Brickenkamp, Schmidt-Atzert & Liepmann, 2010), Befindlichkeit (MDBF-Mehrdimensionaler Befindlichkeitsfragebogen; Steyer, Schwenkmezger, Notz & Eid, 1997) sowie Mind Wandering (Killingsworth & Gilbert, 2010) erfasst. Zusätzlich wurde die mentale und körperliche Anstrengung während der Pausen erhoben (Borg, 1998). Zusätzlich erhoben wurden demographische Informationen (Alter, Gender), Persönlichkeitsmerkmale (BFI-10; Rammstedt, Kemper, Klein, Beierlein & Kovaleva, 2013), Fitnesslevel (1-Item: „Wie fit schätzen Sie sich körperlich ein?), Stresslevel (1-Item: „Wie viel Stress haben Sie derzeit?) und Einstellung zu Pausen (7-Items, z. B. „Eine Pause in Seminaren ist absolut notwendig.“). Die Interventionsphase umfasste 4 Wochen (10. bis 13. Semesterwoche). Die Pre-Post Tests wurden sowohl in der 10. als auch in der 13. Sitzung durchgeführt, um akute (nach nur einer Sitzung) sowie chronische Effekte zu erfassen, was für diese Forschung kennzeichnend ist, aber bisher kaum kombiniert wurde (Ai et al., 2021; Liu et al., 2020).

Ergebnisse

Zur Tagung werden erste Ergebnisse zu akuten Effekten einer Bewegungspause auf die Lernbedingungen Aufmerksamkeit und Befindlichkeit präsentiert. Zwischen Interventions- und Kontrollgruppe wurden vor Beginn der Interventionsphase keine Unterschiede bezüglich Aufmerksamkeit und Befindlichkeit und bezüglich anderer Merkmale (Alter, Gender, Persönlichkeitsmerkmale, Fitnesslevel, Stresslevel, Pausen-Einstellung) festgestellt (alle *p*-Werte > .05). Mittels ANOVAs mit Messwiederholung wurden signifikante Interaktionseffekte (Gruppe*Zeit) bei

zwei von drei Skalen der selektiven Aufmerksamkeit (Skala: Anzahl bearbeitete Zielobjekte: $F(1,100)=11.24$, $p=<.01$, $\eta 2=.142$, Skala: Konzentrationsleistung: $F(1,100)=6.41$, $p=<.01$, $\eta 2=.122$) und bei zwei von drei Skalen zur Befindlichkeit (Skala: Wachheit: $F(1,100)=6,20$, $p=.01$, $\eta 2=.058$; Skala: Innere Ruhe: $F(1,100)= 4.93$, $p=.03$, $\eta 2=.047$) festgestellt. Post-hoc Analysen zeigten, dass sich stets Teilnehmer:innen der Interventionsgruppe über die Zeit mehr als in der Kontrollgruppe verbesserten.

Die Befunde verweisen auf akute positive Effekte einer Bewegungspause auf kognitive und motivationale Lernbedingungen von Studierenden in Hochschulseminaren. Im Vortrag wird auf die weitere Auswertung chronischer Effekte, mögliche Moderationseffekte, Transfermöglichkeiten sowie Limitationen eingegangen.

C-08: Schulentwicklung

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S05 T00 B71
Chair der Sitzung: Marcus Pietsch

The generalizability of instructional leadership and student achievement across contexts: A meta-analytical big data study

Marcus Pietsch[1], Burak Aydin[1], Sedat Gümüs[2]
[1]Leuphana University Lüneburg, Germany; [2]The Education University of Hong Kong, Hong Kong

The relationship between instructional leadership (IL) in schools and student achievement has been debated and empirically studied for decades (Robinson et al., 2008). Although the idea that leaders who attend to the „core business of schools, instruction“ (Urick & Bowers, 2019, p. 251) has been transferred globally, it appears that most studies to date empirically investigating the relation of this kind of leadership to student achievement stem from the Anglophone world (Hallinger et al., 2020). This raises the question of how far it is permissible to generalize such findings on a global scale when they come "from a small and unrepresentative portion of the global population" (Thalmayer et al., 2021, p. 116). This question is crucial as various studies suggest that there are not only significant differences in school leadership across countries and cultures (Bowers, 2020) but also in the relationship between school leadership and student achievement (Karadag, 2020).

We aim to answer this question by adopting a novel methodological approach to dealing with large data sets called Split/Analyze/Meta-Analyze (Cheung & Jak, 2016) to perform meta-analytical integrative and multilevel data analyses that follow comparable analytical protocols (Brunner et al., 2022). Using this approach, we re-analyze raw data from three cycles of the Programme for International Student Assessment (PISA) in which IL were surveyed. Excluding countries with no leadership survey data, the sample includes K1=73, K2=65, K3=70 (or a total of K=75) countries or regions, J1=18162, J2=17549, J3=16400 schools, and N1=515958, N2=480174, N3=519334 students across 2009, 2012, and 2015 cycles respectively. By this means, our study addresses the following research questions:

Research Question 1: What is the overall effect size computed based on the relation between school-level IL scores and the school component of student achievement, and how heterogenous are these effects across countries, cycles, and achievement domains?

Research Question 2: To what extent can observed heterogeneity in the IL-student achievement relation be attributed to cultural and other contextual characteristics?

Hence, to investigate generalizability, we firstly conduct country-specific multilevel analyses and secondly combine results in a three-level meta-analysis. Thirdly, we systematically evaluate the influence of contextual moderators on effect sizes following Cronbach (1982) and, fourthly, validate these results by applying a fully data-driven machine learning approach, which allows identifying interactions between moderators in meta-analyses using a regression and classification tree mechanism (meta-CART; Li et al., 2020).

Regarding our first research question, we find that correlations between IL and student achievement are positive on average in the 2009 PISA cycle (r= 0.033), and negative on average in the two subsequent cycles (r= -0.019 and -0.012). Furthermore, large differences in correlation coefficients can be observed both between and within countries (across measurement occasions). In the meta-

analysis phase, we applyed a three-level random effects meta-analysis. The unconditional three-level model resulted in an average effect size of 0.005 with a lower boundary of -0.021 and an upper boundary of 0.030. Regarding our second research question, we find that these differences are mainly due to the level of human development in a country (HDI) and cultural factors indicated by Hofstede measures. The machine learning approach confirms these results and shows that other contextual factors (i.e. teacher shortage, accountability pressure, school size etc.) are secondary to these influences.

Our study demonstrates that previously reported findings on the relationship between instructional leadership and student achievement should be taken with caution, as it is likely that the results are not generalizable at the global scale. Consequently, policy borrowing is not advisable. What works in one context does not necessarily work in another. Future studies should pay more attention to the survey context and be methodologically more rigorous.

Schulqualität – Unterschiede in den Wahrnehmungen zwischen Schulleitungen und Lehrkräften

Christoph Helm[2], Rolf Strietholt[3], Stephan Gerhard Huber[1]
[1]Pädagogische Hochschule Zug, Schweiz; [2]Johannes Kepler Universität Linz, Österreich; [3]IEA Hamburg

Häufig werden Daten aus Erhebungen im Bildungskontext gesammelt und verwendet, um evidenzbasierte Bildungspolitik zu betreiben. Daten, die für die Entscheidungsfindung im Bildungsbereich (mit hoher Bedeutsamkeit) verwendet werden, sollten zuverlässig und gültig sein (Desimone, 2006; Desimone et al., 2010). Zuverlässigkeit und Gültigkeit können jedoch stark davon abhängen, wer um die Bewertung von Bildungsfragen gebeten wird (z. B. Unterrichtsqualität, Schulqualität). Daher stellt sich schnell die Frage, wie stark sich die Berichte verschiedener Informanten über dasselbe Konstrukt unterscheiden. Wenn Schulleitungen- und Lehrkräfteberichte nicht übereinstimmen, sondern selbst valide Informationen darstellen, stellt sich außerdem die Frage, welche "perspektivenspezifischen Teile" (Kunter & Baumert, 2007) des Konstrukts durch diese beiden Perspektiven repräsentiert werden? Wenn also große Diskrepanzen zwischen zwei oder mehreren Informantengruppen beobachtet werden, stellen sich Fragen danach, welche Information für eine bestimmte Forschungsfrage angemessener/ richtiger ist und wie die Bedeutung perspektivenspezifischer Informationen einzuschätzen ist.

Vor diesem Hintergrund untersuchen wir die Übereinstimmung der Schulleitungs- & Lehrkräfteeinschätzung zur Schulqualität. Es gibt nur wenige Studien, die die Rolle der Perspektiven verschiedener Informanten bei der Bewertung der Schulqualität untersuchen (Bingham et al., 1993; Jacobs, 2018; James et al., 2018; Mitchell et al., 2010; Ramsey et al., 2016; Waasdorp et al., 2011). Bingham et al. (1993) stellten dabei fest, dass Schulleitungen ihre Schulen besser bewerteten als Lehrkräfte. Die Autoren argumentieren, dass diese Unterschiede auf den "Cheerleading-Charakter" der Arbeit von Schulleitungen zurückzuführen sind.

Wir betrachten zwei verschiedene Indikatoren für die Übereinstimmung zwischen Schulleitungen und Lehrkräften. Erstens untersuchen wir den Mittelwertunterschied zwischen den Berichten von Schulleitungen und Lehrkräften über verschiedene Maßnahmen zur Schulqualität. Anhand der Mittelwertunterschiede lässt sich die allgemeine Frage beantworten, inwieweit Schulleitungen und

Lehrkräfte im Allgemeinen übereinstimmen - unabhängig von ihrer Zugehörigkeit zu einer bestimmten Schule. Zweitens untersuchen wir die Korrelationen zwischen den Bewertungen der Schulleitungen und den Bewertungen der Lehrkräfte. Diese Korrelationen berücksichtigen die Zugehörigkeit von Schulleitungen und Lehrkräften zu einer bestimmten Schule und zeigen, ob Lehrkräfte hoch/niedrig bewerten, wenn ihre Schulleitungen dies tun - und umgekehrt.

Methode

Die Daten wurden mittels eines Online-Fragebogens zu Dimensionen der Schulqualität (u.a. Lehrkräftewirksamkeit, Unterrichtskooperation) erhoben, der Teil eines Forschungsprojekts über Schulentwicklungsprozesse in Nordrhein-Westfalen war. Insgesamt nahmen 116 Schulen aller Schularten an der Befragung teil. Für jede der Schulen füllten die Schulleitungen (n = 116, 68 % weiblich) und das Schulpersonal (n = 2380, hauptsächlich Lehrkräfte, 77 % weiblich) den Fragebogen aus. Auf diese Weise erhielten wir hierarchische Daten, die jede Lehrkraft mit ihrer Schulleitung und umgekehrt in Verbindung bringen.

Verwendet wurde der Analyseansatz von (Kim et al., 2018), der ein Cross-Level-Group-Measurement-Invarianz-Modell (CLGMI) vorschlägt, um auf gleiche Messungen auf Lehrkräfte- und Schulleitungsebene zu testen. Das Modell ermöglicht die Schätzung des latenten Mittelwerts der Differenz zwischen der/den Schulleitungsbewertung(en) und der Lehrkräftebewertung. Auf der Grundlage von CLGMI-Modellen verfolgen wir einen zweistufigen Ansatz.

Ergebnisse

Die Angaben von Schulleitungen und Lehrkräften unterscheiden sich bei allen Schulqualitätsdimensionen signifikant: Alle latenten Mittelwerte sind bei den Schulleitungen höher. Allerdings variieren die Übereinstimmungsindizes sowohl bei den latenten Mittelwertunterschieden als auch bei den latenten Korrelationen erheblich.

Unsere Studie zeigt deutlich, dass Schulleitungen alle bewerteten Dimensionen der Qualität ihrer Schulen deutlich positiver einschätzen als ihre Lehrkräfte. Die Effektstärken dieser Beurteilungsunterschiede sind mittelgroß bis groß. Unsere Ergebnisse bestätigen theoretische Annahmen und Konzepte wie den "Cheerleading-Charakter der Schulleitungstätigkeit" und das "Selbst(über)vertrauen der Schulleitungen", die in der Literatur beschrieben werden (Bingham et al., 1993; McCormick, 2001). Unsere Studie stellt daher eine Warnung vor der unreflektierten Interpretation und Verwendung von Daten und Ergebnissen dar, die nur auf einer einzigen Informationsquelle bzw. Perspektive beruhen. Um Daten für die Entscheidungsfindung im Bildungsbereich zu erhalten, die den von Desimone (2006) geforderten Zuverlässigkeits- und Gültigkeitsstandards entsprechen, sind Studien mit mehreren Informanten erforderlich, um valide Daten für die Entscheidungsfindung zu erhalten.

Widerstand im Kontext von Schulentwicklung – lästiges Hindernis oder good sense?

Michelle Bebbon, Svenja Mareike Schmid-Kühn
Universität Koblenz, Germany

Widerstand gilt als konstitutives Moment von Schulentwicklung. Sowohl innerhalb der Schulentwicklungsforschung als auch aus schulpraktischer Perspektive wird die Thematik immer

wieder aufgegriffen. So konstatieren Bonsen und Schratz (2015): „[W]er die Praxis der Schulentwicklung kennt, weiß, dass Veränderungen in der Schule eigentlich immer von Widerstand begleitet werden“ (S.1). Regelmäßig wird das Thema auch von Lehrkräften (z.B. Blume, 2022), Schulleitungen (z.B. Kretschmer, 2017) und der seitens der Schulaufsicht (z.B. DKJS & Stiftung Mercator, 2020) als relevant markiert. Dabei wird Widerstand oftmals im Kontext eines „mentalen Defizitmodell[s]“ verhandelt, das widerstandsleistende Personen als „faul und uneinsichtig diffamiert“ (Philipp und Rolff, 2011, S. 82).

Vor dem Hintergrund dieser Omnipräsenz der Thematik erscheint es umso erstaunlicher, dass Widerstand in der Schulentwicklungsforschung bisher kaum systematisch betrachtet worden ist und die theoretische Konzeptualisierung sowie empirische Erforschung von Widerstand hinter der anderer Disziplinen (vgl. z.B. Johansson & Vinthagen, 2020; Baaz, Lilja & Vinthagen, 2018) zurückbleibt. Huber (2015) konstatiert beispielweise: „Widerstand ist einer der Begriffe, der im Zusammenhang von Schulentwicklung sehr häufig genannt, dabei jedoch immer wieder unklar verwendet wird“ (S. 105).

Diesen Unsicherheiten in Bezug auf die Beschreibung von Widerständen widmet sich der Tagungsbeitrag und leistet, ausgehend von der wissenschaftlichen Begleitung eines Schulentwicklungsprojekts zur Individuellen Förderung an Gymnasien, einen Beitrag zum Aufbau empirisch gesicherten Wissens über Widerstand im Kontext schulischer Entwicklungsprozesse. Dabei wird das Widerstandsphänomen einem interdisziplinären und mehrdimensionalen Ansatz folgend entlang der Fragen nach widerständigen Akteur:innen, Ursachen und Formen von Widerständen als auch Folgen widerständigen Verhaltens sowie dem Umgang mit Widerstand exploriert.

Leitfadengestützten Experteninterviews (vgl. Meuser & Nagel, 2009) mit den betreuenden Koordinator:innen einzelner Fachnetzwerke des Projektes dienen hier als Grundlage einer ersten Beschreibung des Widerstandsphänomens aus Expertensicht. Die Auswertung der Experteninterviews erfolgte theoriegeleitet mittels inhaltlich-strukturierender Inhaltsanalyse (Kuckartz 2018), wobei ein gemischt induktiv-deduktives Vorgehen bei der Kategorienbildung umgesetzt wurde. Als Arbeitsdefinition bzw. sensibilisierendes Konzept (vgl. Kruse, 2016) dient das Verständnis von Widerstand als Erwartungsbruch, wonach Widerstand immer die Negation einer wahrgenommenen Erwartung, z.B. in Form eines (passiven) Unterlassens einerseits oder (aktiven) Gegenwehr andererseits, beinhaltet (vgl. Holzer, 2017, S. 467-68).

Bisherige Ergebnisse zeigen, dass im Kontext von Schulentwicklung nicht, wie oftmals formuliert, nur Lehrkräften die Rolle von Widerständlern zukommt, sondern dass auch Schulleitungen Widerstand leisten. Während Widerstand häufig als Ausdruck persönlicher Bedürfnisse wie z.B. der Abwehr von Mehrbelastung oder der Schaffung eines individuellen Vorteils erscheint, zeigt sich im Projektkontext ebenfalls Widerstand, der die Belange anderer Personen(-gruppen) fokussiert und als *solidarisierender Widerstand* (vgl. Baaz, Lilja, Schulz & Vinthagen, 2016) bezeichnet werden kann. Gleichzeitig richten sich Widerstände nicht nur gegen Veränderungen, sondern sind z.T. auch gegen bestehende Strukturen gerichtet, die von Lehrkräften als hinderlich für den Anstoß von Veränderung erlebt werden.

Im Anschluss an die Forderung von Bennewitz und Thiel (2019) „eine andere Perspektive auf den Untersuchungsgegenstand zu entwickeln, als dies bislang der Fall ist“ (S. 638), wird eine erweiterte Sichtweise auf Widerstand offeriert, die ebenfalls die Frage thematisiert, inwiefern Widerstand nicht nur als lästiges Hindernis, sondern ggf. auch als „good sense“ (Gitlin & Margonis, 1995, S. 379) angesehen werden kann. Im Umgang mit Widerstand werden neben einer gemeinsamen Thematisierung von Problemen innerhalb des Netzwerks und der Möglichkeit einer Kurskorrektur angesichts bestehender Widerstände auch machtbezogene Strategien verfolgt, die durch eine Reduzierung von Freiheiten im Projekt Widerstände zu minimieren versuchen.

Der Ausgangspunkt für Schulentwicklungsprozesse: Die Orientierung der Entwicklungsperspektiven an Stärken, Schwierigkeiten und/oder Herausforderungen

Luisa Grützmacher[1], Julia Holzer[2], Marko Lüftenegger[1,2], Manfred Prenzel[2], Barbara Schober[1]
[1]Zentrum für Lehrer*innenbildung, Universität Wien; [2]Fakultät für Psychologie, Abteilung für Entwicklungs- und Bildungspsychologie, Universität Wien

Theoretischer Hintergrund

Die kontinuierliche Entwicklungen der Gesellschaft macht es notwendig, dass Schulen lernen sich an die veränderten Bedingungen anzupassen (Kools et al., 2020). Solche Lernprozesse sind auch für benachteiligte Schulen von entscheidender Bedeutung, da diese mit besonders herausfordernden Bedingungen konfrontiert sind. Schulen können sich durch fortlaufende intentionale Schulentwicklungsprozesse verändern. Demnach können Schule als lernende Organisationen verstanden werden (Klein, 2017). Die Literatur zu lernenden Organisationen fokussiert stark auf das Lernen durch Misserfolge (Faulstich, 2014; Schechter et al., 2004). Lernen durch Misserfolge muss jedoch nicht zwangsläufig zum Erfolg führen. Außerdem werden wichtige Lernmöglichkeiten ausgeschlossen. Auch Herausforderung und Erfolge können wichtige Impulse für Lernprozesse darstellen (Karsten et al., 2000; Schechter et al., 2004).

Forschungsfragen

Diese Studie zielt darauf ab, ein besseres Verständnis über die Situation von benachteiligten Schulen sowie deren Orientierung im Schulentwicklungsprozess zu gewinnen. Demnach werden Stärken, Schwierigkeiten, Herausforderungen und Entwicklungsperspektiven von benachteiligten Schulen untersucht. Aus den Übereinstimmungen der Entwicklungsperspektiven mit den berichteten Stärken, Schwierigkeiten oder Herausforderungen kann abgeleitet werden, ob Schulen im Schulentwicklungsprozess versuchen vorhandene Stärken auszubauen, Schwierigkeiten abzubauen oder Herausforderungen zu bewältigen. Keine Übereinstimmung deutet darauf hin, dass sich die Schulen an nicht identifizierten Faktoren orientieren. Die Forschungsfragen lauten:

1. Welche (a) Stärken, (b) Schwierigkeiten, (c) Herausforderungen und (d) Entwicklungsperspektiven werden von Schulleitungen von Volks- und Mittelschulen angegeben?
2. Welche Profile liegen der Orientierung der Entwicklungsperspektiven auf Stärken, Schwierigkeiten, Herausforderungen oder anderen Faktoren in (a) Schulen und (b) Volks- und Mittelschulen zugrunde?

Methode

Die Stichprobe umfasst 100 benachteiligte Schulen in Österreich (55 Volksschulen/ 45 Mittelschulen). Die Einstufung als benachteiligte Schule erfolgte durch einen Index, der die sozioökonomische Zusammensetzung der Schüler*innenschaft abbildet. Die Schulleitungen wurden gebeten einen

Online-Fragebogen mit offenen Fragen zu Stärken, Schwierigkeiten, Herausforderungen und Entwicklungsperspektiven auszufüllen.

In dieser Studie wurde ein Mixed-Methods-Ansatz eingesetzt. Die qualitativen Daten wurden mittels qualitativer Inhaltsanalyse (Mayring, 2012) analysiert. Dazu wurde ein Kategoriensystem mit zwei Dimensionen entwickelt (Handlungsebenen/ Qualitätsdimensionen). Zwei Forscher*innen haben das Kategoriensystem unabhängig voneinander auf die Daten angewandt. Die vergebenen Kategorienlabels wurden gegengeprüft (gute Übereinstimmung, $\omega = .88$), die Daten für quantitative Analysen aufbereitet und Häufigkeiten zur Beantwortung der ersten Forschungsfrage in SPSS berechnet. Zur Beantwortung der zweiten Forschungsfrage wurden in Mplus latente Profilanalysen durchgeführt. Die Anzahl der Profile wurde durch den BIC und Entropie-Index bestimmt.

Ergebnisse & ihre Bedeutung

Die Ergebnisse zur ersten Forschungsfrage zeigen die vielfältigen Herausforderungen, Schwierigkeiten, Stärken und Entwicklungsperspektiven der Schulen. Während einige Aspekte selten genannt wurden, gab es bei anderen Aspekten eine hohe Übereinstimmung. Viele Schulen berichten beispielsweise, dass ihnen die Gestaltung einer positiven Schulkultur gut gelingt (76 Schulen), ihnen die Kooperation mit Familien und Elternarbeit Schwierigkeiten bereitet (68 Schulen), sie den Lehrpersonenmangel als Herausforderung erleben (68 Schulen) und sie die Unterrichtsgestaltung weiterentwickeln wollen (67 Schulen). Volks- und Mittelschulen geben ähnliche Stärken, Schwierigkeiten, Herausforderungen und Entwicklungsperspektiven an. Es gibt allerdings auch Unterschiede. So geben beispielsweise Mittelschulen häufiger an, dass sie ihre Stärke in den non-formalen Strukturen der Schule sehen (Δ18 %).

Die Ergebnisse zur zweiten Forschungsfrage zeigen, dass es unterschiedliche Profile bei der Orientierung der Entwicklungsperspektiven gibt. Während sich die Entwicklungsperspektiven einer Gruppe von Schulen eher an den Schwierigkeiten und anderen Faktoren orientieren, orientiert sich eine weitere Gruppe eher an den Stärken und anderen Faktoren. Eine dritte Gruppe weist eine stärkere Orientierung an anderen Faktoren aus, während die Orientierung an Stärken, Schwierigkeiten und Herausforderungen eher gering und ausgewogen ist. Schulen mit ähnlichen herausfordernden Bedingungen haben demnach unterschiedliche Ansätze im Schulentwicklungsprozess. Der Vergleich zwischen Volks- und Mittelschulen zeigt, dass sich die Schulformen in der Anzahl und Ausprägung der identifizierten Profile unterscheiden.

Insgesamt ergibt sich ein komplexes Bild zur Situation und Schulentwicklungsorientierung von benachteiligten Schulen, die keine Verallgemeinerung zulässt. Die Bedeutung der nicht identifizierten Faktoren soll im Beitrag diskutiert werden.

C-09: Digitales Lernen II

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S05 T00 B83
Chair der Sitzung: Simon Schultze

Spieler:innentypen in digitalen Lernsysteme - Erste Ergebnisse zur Gamification User Types Hexad Skala

Simon Schultze, Jan Hochweber
Pädagogische Hochschule St. Gallen, Switzerland

Theorie

Frühere Untersuchungen zeigen, dass Gamification-Elemente (GE) die Motivation und das Engagement von Schüler:innen steigern können (z. B. Seufert et al., 2017). Allerdings ist die Wirkung stark abhängig vom zugrunde liegenden Gamification-Konzept (Araya et al., 2019). Darin müssen GE bestimmt bzw. entworfen werden, die zu den jeweiligen Zielen und dem Nutzungskontext passen. Dabei ist zu beachten, dass verschiedene Nutzer:innen unterschiedliche GE bevorzugen (Hamari et al., 2014). Als Erklärung für diese individuellen Präferenzen wurde die Unterscheidung zwischen verschiedenen Spieler:innentypen vorgeschlagen. Das bekannteste Spieler:innentypenmodell für Gamification ist das Gamification User Types Hexad-Modell (Marczemski, 2015). Ein auf diesem Modell basierendes Instrument, die Hexad-Skala, wurde von Tondello et al. (2016, 2019) entwickelt. Seine psychometrischen Eigenschaften wurden in ersten Studien empirisch untersucht, bislang jedoch nicht für den Kontext digitaler Lernsysteme.

Fragestellung

In dieser Studie leisten wir einen Beitrag zur empirischen Forschung über das Hexad-Modell sowie zur Validierung der Hexad-Skala, indem wir:

a) …die Faktorenstruktur der Hexad-Skala mithilfe einer konfirmatorischen Faktorenanalyse (CFA) prüfen,

b) …die Unterscheidung verschiedener Spieler:innenprofile auf der Grundlage einer latenten Profilanalyse (LPA) mit den Hexad-Faktoren untersuchen.

Methode

Teilnehmer:innen

Im Kontext einer Nutzungsstudie für das digitale Lernsystems Lernnavi nahmen 318 Schüler:innen aus 18 Klassen der gymnasialen Oberstufe (Sekundarstufe II) in der Deutschschweiz und Liechtenstein an einer Online-Befragung teil. Zur Rekrutierung der Stichprobe wurden Mathematiklehrer:innen mit ihrer gesamten Klasse zur Befragung eingeladen, mit der Intention, so die Teilnahme von Schüler:innen mit unterschiedlichen soziodemographischen Hintergründen und verschiedenen Leistungsniveaus sicherzustellen.

Erhebung

Alle Variablen wurden mit einem Online-Fragebogen erhoben.

Spieler:innentypen: Für die Spieler:innentypen wurde der Fragebogen von Tondello et al. (2016, 2019) ins Deutsche übersetzt. Er enthält 24 Items, die erheben, wie und warum Personen spielen (z. B. "Ich mag es, Teil eines Teams zu sein"). Die 24 Items sind sechs Hexad-Skalen zugeordnet, die sechs verschiedene Spieler:innentypen (z. B. Achiever) repräsentieren.

Soziodemographische Merkmale: Daneben wurden soziodemographische Merkmale der Schüler:innen erhoben (wie Geschlecht, Alter und SES der Familie).

Vorgehen

Die Schüler:innen erhielten den Online-Fragebogen 2-3 Wochen nach Beginn der Nutzungsstudie zu Lernnavi im Frühjahr 2022. Von den 345 eingeladenen Schüler:innen füllten 318 den Fragebogen aus, was einer Rücklaufquote von 92% entspricht.

Ergebnisse

Die Faktorenstruktur der Hexad-Skalen wurde mithilfe einer mehrfaktoriellen CFA in Mplus überprüft. Für jede der sechs Dimensionen wurde (analog zum Vorgehen bei Tondello et al., 2019) ein eigener Faktor spezifiziert. Gängige Kriterien der Modellpassung deuten auf eine (noch) akzeptable Passung des Modells hin (CFI = 0.891, RMSEA = 0.070), die insgesamt günstiger ausfällt als bei Tondello et al. (2019). Für die zweite Fragestellung wurde eine Latente Profilanalysen (LPA) berechnet. Als statistisch angemessen (auf Basis von Informationsindizes, der Entropie sowie Likelihood Ratio Tests; vgl. Spurk et al., 2020) sowie inhaltlich gut interpretierbar erwies sich eine 4-Profile-Lösung. Auffällig ist, dass die vier Profile nicht gut mit den Spieler:innentypen des Hexad-Modells in Übereinstimmung zu bringen sind. So ist z.B. eine Profilgruppe dadurch gekennzeichnet, dass sie die abgefragten GE generell eher ablehnt, während die grösste Profil-Gruppe (47.2%) über alle GE hinweg hohe Zustimmungswerte aufweist – ausser bei den vom Spieler:innentyp «Disruptor» im Hexad-Modell präferierten GE, welche bei allen Profil-Gruppen durchfallen (z.B. die Möglichkeit bestimmte Spielmechaniken zu überlisten).

Bislang haben sich Studien zum Hexad-Spieler:innentypenmodell auf Gamification unabhängig vom Nutzungskontext fokussiert. Im Gegensatz dazu liefert der vorliegende Beitrag Erkenntnisse speziell für den Kontext digitaler Lernsysteme und für eine klar definierte Schüler:innenpopulation, die bisher nicht untersucht wurde. Wichtige Limitationen des Modells bzw. der Analysen sind, dass sich aus CFA wie auch LPA Hinweise auf mögliche Anpassungsbedarfe am Hexad-Modell ergeben, und v.a. die Ergebnisse der LPA einer Bestätigung mit einer grösseren Stichprobe von Schüler:innen bedürfen.

Die Bedeutung digitaler Medien für die unterrichtsbezogene Lehrerkooperation – erste Ergebnisse einer empirischen Studie

Mario Engemann, Bardo Herzig
Universität Paderborn, Deutschland

Die Herausforderungen des Digitalisierungsprozesses als Mediatisierungsschub tangieren das öffentliche und private Leben. Anfangs als ein rein technischer Prozess zur Wandlung analoger in digitale Signale betrachtet, steht Digitalisierung inzwischen kennzeichnend für die Omnipräsenz digitaler Medien im Alltag (vgl. Herzig & Martin 2018). Davon sind selbstredend auch die Schule und der Unterricht betroffen, die sich ohne eine Berücksichtigung der Digitalisierung nicht mehr angemessen (re-)konstruieren lassen. An die Schulen wird die Aufgabe herangetragen, die Herausforderungen nicht nur aufzugreifen, sondern den durch Medien geprägten Wandel konstruktiv für sich mitzugestalten. Der Digitalisierungsprozess von Schule und Unterricht bedarf zugleich einer

kontinuierlichen Professionalisierung der Lehrkräfte (vgl. Engemann & Herzig 2021). In diesem Kontext gilt die unterrichtsbezogene Lehrerkooperation „als Königsweg der Qualitätssteigerung und Professionalität“ (Terhart 2001) – wird in der Literatur doch mit ihr der Anspruch verbunden, unterschiedliche schulische Probleme zu lösen (vgl. Drossel 2015). In Bezug auf digitalisierungsbezogene Lehrerkooperationen ist die Forschung in summa rar, entsprechend die empirische Datenbasis überschaubar (vgl. Engemann & Herzig 2021; Heldt et al. 2020). Insbesondere fehlen Arbeiten zu einzelnen Schulfächern, was in dieser Studie aufgegriffen wird. Zielgruppe des Projekts sind Lehrkräfte des gymnasialen Schulfaches Erziehungswissenschaft in Nordrhein-Westfalen. Es liegt dabei die Annahme zugrunde, dass die Digitalisierung für ‚kleinere‘ Fachschaften Potenziale hinsichtlich einer erweiterten fachbezogenen Kooperation bietet (vgl. ebd.). Ausgehend von dieser unzureichenden Forschungslage zur digitalisierungsbezogenen Lehrerkooperation setzt sich das Projekt das Ziel, den Stellenwert digital-kooperativer Arbeitsbeziehungen von Lehrkräften – exemplarisch am gymnasialen Schulfach Erziehungswissenschaft – herauszuarbeiten. Die zentrale Fragestellung der Studie lautet: Wie stellt sich die Lehrerkooperation innerhalb des gymnasialen Schulfaches Erziehungswissenschaft unter Berücksichtigung der Digitalisierung in Formen, Gestaltung und Rahmenbedingungen dar? Dabei lässt sich in Bezug auf die Digitalisierung die unterrichtsbezogene Lehrerkooperation auf drei Ebenen ausgestalten: 1.) auf der Prozessebene, verstanden als Kooperation via digitale Medien, 2.) auf der Ebene der inhaltlichen Kooperation mit digitalen Medien als Gegenstand und 3.) bezogen auf den mediendidaktischen Einsatz im Unterricht, über den sich Lehrkräfte austauschen. Das dieser Studie zugrundeliegende Kooperationsverständnis orientiert sich an der organisationspsychologischen Definition von Spieß (2004), wonach „Kooperation […] durch den Bezug auf andere, auf gemeinsame zu erreichende Ziele bzw. Aufgaben [gekennzeichnet ist], sie ist intentional, kommunikativ und bedarf des Vertrauens. Sie setzt […] Autonomie voraus und ist der Norm der Reziprozität verpflichtet“ (Spieß 2004). Als Referenz für Formen von unterrichtsbezogener Lehrerkooperation dient das organisationspsychologische Konzept der Wuppertaler Forschungsgruppe um Gräsel (2006), in dem die Formen *Austausch*, *Synchronisation* und *Ko-Konstruktion* unterschieden werden (vgl. Gräsel et al. 2006). Methodisch wurde unter Berücksichtigung bestehender Items zur Lehrerkooperation (vgl. Drossel et al. 2020; Fussangel 2008) ein quantitatives Erhebungsinstrument konzipiert. Ausgangspunkt für den konzeptionierten Fragebogen ist der Ansatz zur Differenzierung von Kooperationsformen nach Gräsel et al. (2006). Im Online-Fragebogen wurden Aspekte der Digitalisierung, des Schulfaches Erziehungswissenschaft und die Folgen der Coronapandemie berücksichtigt. Die Datenerhebung fand von Dezember 2021 bis Februar 2022 statt. Die Ergebnisse beruhen auf der Auswertung von 282 Lehrkräften des gymnasialen Schulfaches Erziehungswissenschaft. Als Ergebnisse der deskriptiven und interferenzstatistischen Verfahren zeigen sich beispielsweise signifikante Unterschiede hinsichtlich der untersuchten Schulformen des allgemeinbildenden Gymnasiums, des beruflichen Gymnasiums und Gesamtschule. Am beruflichen Gymnasium konnte festgestellt werden, dass schulintern distanziertere Einstellungen zur Lehrerkooperation vorliegen, sich diese jedoch nicht in den Kooperationspraktiken widerspiegeln. Vielmehr suchen Lehrkräfte des Schulfaches Erziehungswissenschaft am beruflichen Gymnasium die

schulübergreifende Zusammenarbeit mit ihren Fachkolleg*innen. Zugleich zeigt sich allgemein, dass soziale Medien bzw. digitale Kommunikationstools – mit Ausnahme von E-Mails und Messenger-Diensten wie WhatsApp – für die Kooperation nur eine untergeordnete Rolle einnehmen. Im Vortrag werden die theoretischen Perspektiven, die Konzeption des Projekts, das Erhebungsinstrument sowie ausgewählte Ergebnisse vorgestellt. Die quantitativen Ergebnisse bilden die Grundlage für eine weitere Analyse mit qualitativen Verfahren, die im Anschluss durchgeführt wird.

Warum nutzen Lehrkräfte Augmented Reality (nicht)? - Eine multimethodische Untersuchung von Gelingensfaktoren

Kristin Altmeyer, Sarah Malone, Luisa Lauer, Markus Peschel, Roland Brünken
Universität des Saarlandes, Saarbrücken, Deutschland

Theoretischer Hintergrund

Durch Augmented Reality (AR) können z.B. mithilfe eines Tablets oder Smartglasses virtuelle Objekte in eine reale Umgebung integriert werden (Azuma et al., 2001), um so Unsichtbares sichtbar zu machen und 3D-Informationen zu visualisieren. Evidenz aus Metaanalysen unterstreicht das lernförderliche Potenzial von AR (Garzón & Acevedo, 2019). Trotz empirisch nachgewiesener Lernvorteile und obwohl Lehrpersonen ihr Interesse an AR bekunden, findet AR bisher wenig praktische Anwendung im Bildungswesen (Tzima et al., 2019).

Das Will-Skill-Tool Modell (Knezek & Christensen, 2016; Knezek et al., 2003) versucht erfolgreiche Technologieintegration in den Unterricht zu erklären. Der Faktor Will meint die Einstellung von Lehrpersonen zur Nutzung von Technologie im Unterricht. Skill beschreibt die technologischen Fähigkeiten von Lehrpersonen und Tool beschreibt den Zugang zu Technologien an einer Bildungseinrichtung. Während in vorausgehenden Studien durch diese Faktoren mindestens 60% der Varianz des Einsatzes digitaler Medien aufgeklärt werden konnten (Petko, 2012) bleibt unklar, inwiefern die Vorhersagekraft des Modells auf den spezifischen Einsatz von AR im Unterricht übertragbar ist.

Fragestellung

Nach Tzima (2019) ist die Rolle der Lehrpersonen für die erfolgreiche Integration von AR in den Unterricht wesentlich. Das Ziel dieser Arbeit ist es daher, auf Grundlage von Lehrpersonenbefragungen den Einsatz von AR im Bildungskontext zu untersuchen und Gelingensbedingungen zu identifizieren.

Dazu wird im Rahmen von Forschungsfrage 1 quantitativ untersucht, wie die Faktoren Will, Skill und Tool zur spezifischen Erklärung von AR-Integration in den Unterricht beitragen.

Im Rahmen von Forschungsfrage 2 wird qualitativ untersucht, welche Potenziale und Herausforderungen AR-erfahrene Lehrpersonen dem Einsatz von AR im Unterricht zuschreiben.

Methode

Im Sinne eines multimethodischen Ansatzes wurden zwei parallele Teilstudien durchgeführt. Bei der ersten Teilstudie handelte es sich um eine quantitative Befragung von Lehrpersonen mithilfe eines Onlinefragebogens. N=30 Lehrkräfte (MAlter=41.04, SDAlter=9.14, 75% weiblich) beantworteten Items zu den Faktoren Will, Skill und Tool. Will wurde mit einer für AR adaptierten Version des

Fragebogens zur Werteüberzeugung zur Nutzung digitaler Medien im Unterricht erfasst (Rubach & Lazarides, 2021). Skill wurde mithilfe des für AR angepassten Fragebogens für Kompetenzüberzeugungen im Bereich digitaler Medien für Lehrkräfte erhoben (Rubach & Lazarides, 2021). Für den Faktor Tool wurde erfragt, welchem Anteil der Lehrer- bzw. Schülerschaft der eigenen Bildungseinrichtung verschiedene (AR-)Endgeräte als individuelle Arbeitsmittel zur Verfügung gestellt werden können. Außerdem wurde mithilfe der für AR angepassten stages of adoption scale (Knezek et al., 2003) erfasst, inwiefern befragte Lehrpersonen bisher selbst AR in ihrem Unterricht eingesetzt haben.

In der zweiten, qualitativen Teilstudie nahmen n=10 Lehrkräfte im Alter zwischen 28 und 58 Jahren (20% weiblich), die bereits AR im Unterricht nutzen, an einem semi-strukturierten Interview teil. Dabei beantworteten sie Fragen zu Gelingensbedingungen, Potenzialen und Herausforderungen des Einsatzes von AR im Unterricht. Die Fragen bezogen sich auf die drei Phasen Erstellung von AR-Inhalten, Vorbereitung des Unterrichtes mit AR und Anwendung von AR. Nach Transkription der Interviews wurde die qualitative Inhaltsanalyse nach Mayring (2022) angewandt.

Ergebnisse

Eine Regressionsanalyse zur Beantwortung von Forschungsfrage 1 zeigte, dass durch die Faktoren Will, Skill und Tool gemeinsam R^2=43% der Varianz des unterrichtlichen Einsatzes von AR aufgeklärt werden konnten (p=.071). Die Faktoren konnten einzeln jedoch nicht signifikant Varianz aufklären. Die spezifische Erklärung von AR-Integration scheint somit komplexer und vielschichtiger zu sein als die Erklärung des generellen Technologieeinsatzes im Unterricht.

Bezüglich Forschungsfrage 2 zeigte sich durch Analyse der Interviewantworten, dass die größte Herausforderung beim Erstellen individueller AR-Inhalte das Fehlen zuverlässiger und beständiger Apps darstellt. Bestehende AR-Inhalte sind dagegen häufig nicht auf Unterrichtsziele abgestimmt oder weisen inhaltliche Schwächen auf. Die Mehrheit der interviewten Lehrpersonen empfindet Fortbildungen und mehr Vorbereitungszeit als notwendig, um AR gewinnbringend im Unterricht einzusetzen. Die meisten Lehrkräfte sehen großes Potenzial in der Möglichkeit, durch AR 3D-Inhalte zu visualisieren und schreiben AR eine sehr hohe Relevanz für zukünftigen Unterricht zu.

Tell Show Enact Do: Ein Lerndesign zur Förderung des Einsatzes von Augmented und Virtual Reality im Unterricht

Josef Buchner
St. Gallen University of Teacher Education, Schweiz

1 Theoretischer Hintergrund

Augmented Reality (AR) und Virtual Reality (VR) Technologien werden vielfältige Potentiale für die Gestaltung von Lernszenarien zugeschrieben; z.B. das Sichtbar-machen nicht-sichtbarer Phänomene oder das Trainieren komplexer Handlungen (Akçayır & Akçayır, 2017; Zender et al., 2020).

Die Einlösung dieser Potentiale im Unterricht hat sich jedoch als herausfordernd für Lehrpersonen herausgestellt. Zum Beispiel fehlen konkrete Umsetzungsbeispiele für die Praxis sowie die Fähigkeit, AR/VR-basierte Lernumgebungen zu gestalten (Alalwan et al., 2020; Buchner et al., 2022). In der

Lehrpersonenbildung muss diese Problematik berücksichtigt werden und der Einsatz von AR/VR vorbereitet werden.

Buchner und Hofmann (2022) haben dafür das auf dem SQD-Modell (Tondeur et al., 2012, 2021) basierte *Tell-Show-Enact-Do* (TSED) Lerndesign entwickelt. Dieses sieht vor, dass Lehrpersonen theoretisches und praktisches Wissen über den Einsatz von AR/VR erhalten (*Tell, Show)*, authentische Erfahrungen mit AR/VR-Lernszenarien machen (*Enact)* und darauf aufbauend eigene AR/VR-Unterrichtsszenarien entwickeln (*Do*). In Buchner und Hofmann (2022) wurde das TSED Lerndesign mit einem Lerndesign verglichen, dass weniger stark auf dem SQD-Modell basiert. Die Ergebnisse zeigen, dass das TSED Lerndesign drei Faktoren der Technologieintegration, *Will, Skill, Tool* (Knezek et al., 2003; Knezek & Christensen, 2016), signifikant stärker fördert als das in der Kontrollgruppe eingesetzte TSE (*Tell, Show, Enact*) Lerndesign.

Eine wesentliche Limitation dieser Studie war, dass in der Kontrollgruppe die *Do*-Phase gänzlich fehlte. Die Bedeutsamkeit dieser Phase hat sich jedoch in anderen Studien gezeigt (e.g. Hsu & Lin, 2020). Die Adressierung dieser Limitation ist das Ziel der vorliegenden Studie.

2 Fragestellung

In dieser Studie wird untersucht, ob das gänzlich auf dem SQD-Modell basierte TSED Lerndesign effektiver für die Förderung AR/VR-spezifischer Faktoren der Technologieintegration (*Will, Skill, Tool*) ist als ein Lerndesign, das zwar die *Do*-Phase integriert, jedoch auf die *Enact*-Phase verzichtet (TSD Lerndesign).

Angelehnt an Tondeur et al. (2012) sowie Buchner und Hofmann (2022) wird vorhergesagt, dass das TSED Lerndesign besser in der Lage ist, diese Faktoren zu fördern als das TSD Lerndesign.

3 Methode

Um die Hypothesen zu testen, wurden die beiden Lerndesigns in einer jährlich durchgeführten Online-Lehrveranstaltung implementiert. Für diese Studie werden Daten aus den Jahren 2020 (n = 29) und 2022 (n = 20) herangezogen. Welches Lerndesign in welchem Jahr implementiert wurde, erfolgte zufällig.

Um den Einfluss des jeweiligen Lerndesigns auf die Faktoren *Will, Skill, Tool* zu messen, wurde auf die von Buchner und Hofmann (2022) adaptierte TUI-Skala (Kothgassner et al., 2013) zurückgegriffen (*Will*, 6 Items, α = 0,76; *Skill,* 3 Items, α = 0,71; *Tool,* 3 Items, α = 0,73).

4 Ergebnisse und Relevanz

Die Ergebnisse des t-Test für unabhängige Stichproben zeigen, dass die Lehrpersonen in der TSED-Bedingung signifikant höhere Werte für *Will* ($t(47) = -1{,}69$, $p = 0{,}049$, einseitige Testung, $d = 0{,}49$), *Skill* ($t(47) = -1{,}82$, $p = 0{,}038$, einseitige Testung, $d = 0{,}53$), und *Tool* ($t(47) = -3{,}25$, $p = 0{,}002$, $d = 0{,}95$) berichten als die Lehrpersonen in der TSD-Bedingung.

Die Ergebnisse bestätigen, dass das ganzheitlich auf dem SQD-Modell basierte TSED Lerndesign effektiv ist, um Lehrpersonen auf den Einsatz von AR/VR im Unterricht vorzubereiten. Zudem wird bestätigt, dass die Empfehlungen des SQD-Modells für die Lehrpersonenbildung relevant sind, um auf die Integration von Technologien vorzubereiten.

Die Studie erweitert die in Buchner und Hofmann (2022) berichtete Limitation der fehlenden *Do*-Phase und kann zeigen, dass die *Enact*-Phase ein bedeutsamer Bestandteil des Lerndesigns ist.

Eine Limitation der Studie ist, dass keine Daten hinsichtlich der Nutzung von AR/VR in der Unterrichtspraxis berichtet werden können. Solche Daten sollten in zukünftigen Studien erhoben werden.

Das TSED Lerndesign hat sich in dieser Studie erneut als effektiv für die Vorbereitung der Technologieintegration im Unterricht herausgestellt. Entsprechend kann das Lerndesign für die Nutzung in der Lehrpersonenbildungspraxis empfohlen werden.

C-10: Lehrkräfte und e-Learning

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S05 R03 H20
Chair der Sitzung: Anika Kohmer

Comparison of trainee teachers‘ and business education students’ ability to deal critically with online information

Marie-Theres Nagel, Anika Kohmer, Olga Zlatkin-Troitschanskaia, Lisa Martin de los Santos, Jennifer Fischer, Andreas Maur
Johannes Gutenberg-Universität, Germany

Objectives

Nowadays, the Internet has become an indispensable part of teachers' everyday work, both in the course of using digital media for lesson planning and implementation, and for teaching digital skills in the classroom [1,2,3,4,5]. In particular, aspiring teachers need to rely on (digital) information research, as they have not yet built up their own collection of materials that they can draw on for lesson planning [6]. However, greater Internet use does not automatically go hand in hand with greater competence in using online media, but competently using online information, hereafter described as “critical online reasoning” (COR; for details, [7]), also poses a challenge for graduates and prospective teachers ([8]; [9]). Therefore, systematic fostering these competencies in teacher education is important to properly prepare them for their daily, professional work [10,11]. Despite the necessity, however, the use of digital media has not yet been systematically curricularly anchored and is only rarely an explicit component of teacher education [12]. This is also confirmed by studies in which trainee teachers show substantial competence deficits, e.g., to critically evaluate online information [13,14]. At the same time, there is a lack of studies on how these competencies and their different facets develop in students in teacher education [15]. This research question is to be addressed with the here presented analyses.

Theory/Hypotheses

The definition of the COR construct describes personal skills in searching, selecting, accessing, processing, and using online information to solve a given (information-based) problem or build knowledge while critically distinguishing high-quality from low-quality information and reasoning argumentatively with credible and relevant online information. COR comprises three interconnecting facets: (i) Online Information Acquisition, (ii) Critical Information Evaluation, and (iii) Reasoning based on Evidence, Argumentation, and Synthesis [7]. The contribution presents recent findings on how the COR ability is developed among students of business education in comparison to trainee teachers of business and economics and to what extent differences in the types of sources used are observable.

Method

Based on COR’s conceptual-theoretical framework, we developed the new COR Assessment (CORA), which measures the use of online information in an authentic online environment while gathering process data (log files with time stamps) and performance data (written answers). Participants are asked to complete a computer-based performance task in which they are given 20 minutes to perform an open web search on an initial question, judge the online information they use to answer the question, and write an open-ended response.

Subsequently, the responses are evaluated by independent human raters based on a rating scheme, which was developed based on the COR construct definition [7] and the log files are analyzed regarding the number of online sources as well as the quality and type of the websites used while answering the CORA task. The online sources are classified regarding their quality with a categorization scheme that was developed based on established research approaches [16,17,18].

Results

In the 2021 summer term, 46 trainee teachers and 87 students in the bachelor's program in business education participated in CORA (for a sample description, see Table 1). Overall, both groups perform almost equally well (trainee teachers *m*=3.15, students *m*=3.15), although the results vary significantly depending on the competency facet considered (for example, in the facet "weighing arguments": trainee teachers *m*=2.96, students *m*=2.47).

The findings provide unique insights into which skill requirements young professionals already fulfil to competently use online information and which COR facets need more targeted fostering in professional training and education. From the results presented, implications for the training of such competencies in teacher education are discussed.

Förderung des Selbstregulierten Lernens bei Lehramtsstudierenden durch E-Learning

Nathalie Barz, Manuela Benick, Laura Dörrenbächer-Ulrich, Franziska Perels
Universität des Saarlandes, Deutschland

Lernen mit digitalen Medien gewinnt im Bildungskontext immer mehr an Bedeutung. So ermöglichen digitale Lernumgebungen den Erwerb und die Unterstützung von Fähigkeiten, die dem lebenslangen Lernen dienen, wie z.B. Selbstreguliertes Lernen (SRL), welches positiv mit der akademischen (Richardson et al., 2012) und beruflichen Leistung (Sitzmann & Ely, 2011) korreliert.

Zimmerman (2011) definiert SRL als Prozess, bei dem Lernende Kognitionen, Emotionen und Verhalten aktivieren und aufrechterhalten, um persönliche Ziele zu erreichen. In seinem sozial-kognitiven Modell unterteilt Zimmerman (2000) SRL in drei Phasen, welche jeweils kognitive, metakognitive und motivationale Strategien umfassen. In der Planungsphase setzen sich die Lernenden Ziele und wählen Lernstrategien basierend auf ihrem Vorwissen aus (Perels et al., 2020). In der Handlungsphase werden kognitive und motivationale Strategien angewendet, um Wissen zu konsolidieren (Artelt & Wirth, 2014) und die Lerntätigkeit aufrecht zu erhalten (Baumann & Kuhl, 2013). Zudem wird der Lernprozess überwacht, um ihn ggf. anzupassen. In der Reflexionsphase evaluieren die Lernenden ihre Lernziele in Bezug auf die Zielerreichung. Diese Selbstreflexion hat wiederum Auswirkungen auf den folgenden Lernprozess. Vor allem in der Ausbildung von Lehrkräften ist SRL besonders bedeutsam, weil Lehramtsstudierende, die ihr Wissen und ihre Kenntnisse bezüglich SRL verbessern, diese einerseits explizit weitergeben können und andererseits implizit als Vorbilder für ihre zukünftigen Schüler*innen fungieren können (Peeters et al., 2014).

Forschungsergebnisse zeigen, dass angehende Lehrkräfte über ein geringes (Glogger-Frey et al., 2018) oder fragmentiertes Wissen zu SRL-Lernstrategien verfügen (van Deur et al., 2016) und wenige SRL-Strategien anwenden, um ihr eigenes Lernen zu steuern (Lawson et al., 2019).

Bezogen auf digitale Lernumgebungen gibt es jedoch nur begrenzte empirische Evidenz dazu, wie SRL gefördert werden kann. Obwohl bisherige Studien im Zusammenhang mit der Förderung von SRL durch E-Learning bei Studierenden vielversprechende Ergebnisse zeigen (Alkhasawnh & Alqahtani, 2019; Bellhäuser, 2016), bedarf es für die Zielgruppe der Lehramtsstudierenden weiterer Forschung.

Zielsetzung der vorliegenden Studie war es daher, den Effekt eines digitalen SRL-Trainings für angehende Lehrkräfte zu untersuchen. Dabei wurde angenommen, dass Teilnehmende, die ein SRL-Training erhielten, anschließend mehr SRL-Strategien einsetzen werden und mehr SRL-Wissen im Posttest abrufen können als Teilnehmende in der aktiven Kontrollgruppe.

Der Studie lag ein Prä-Post-Kontrollgruppen-Design zugrunde: Jeweils eine Befragung wurde zu Beginn und nach dem sechswöchigen Training durchgeführt. Zwischen den Messungen erwarben die Teilnehmenden der E-Learning Bedingung über die Online-Plattform „*Moodle*“ Wissen und Kompetenzen über SRL und SRL-Strategien, während die aktive Kontrollgruppe ein Seminar ohne SRL-Bezug besuchte.

Die Stichprobe umfasste N = 158 Lehramtsstudierende (M_{Alter} = 24.50, *SD* = 4.18, 70.9% weiblich) einer deutschen Universität, die quasi-randomisiert der E-Learning- (n = 76) oder der Kontrollbedingung zugeordnet wurden (n = 82). Die SRL-Intervention umfasste sechs Module basierend auf Zimmermans SRL-Modell (2000). Jedes Modul umfasste theoretische Informationen zu SRL-Strategien sowie digitale Arbeitsblätter und Selbsteinschätzungen. Die Intervention basierte auf einem bereits positiv evaluierten Training (anonymisiert). Die Verwendung von SRL-Strategien wurde anhand eines SRL-Fragebogens mit 48 Items mit dreizehn Subskalen gemessen (anonymisiert) mit zufriedenstellenden Reliabilitäten für die Gesamtskala (Cronbachs = .93 (Prätest), und = .97 (Posttest) sowie die Subskalen (Cronbachs = .58 - .90). Das SRL-Wissen wurde anhand eines SRL-Wissenstests, bestehend aus zehn offenen Fragen, gemessen (anonymisiert).

Die Ergebnisse der (M)ANOVA mit Messwiederholung zeigen eine signifikante Interaktion Zeit x Gruppe, mit einem Zuwachs an SRL-Wissen nach dem Training in der E-Learning-Bedingung ($F(1,119) = 56.16$, $p < .001$). Eine signifikante Interaktion bezüglich der Verwendung von SRL-Strategien ($F(1,119) = 0.52$, $p = .473$) konnte nicht festgestellt werden. Die Analyse auf Subskalen-Niveau hinsichtlich der SRL-Phasen und -komponenten führte zu ähnlichen Ergebnissen.

Die Ergebnisse unterstreichen die Effektivität von digitalen SRL-Trainings für Lehramtsstudierende hinsichtlich des Wissenserwerbs, aber um den tatsächlichen SRL-Strategieeinsatz von Lehramtsstudierenden mit Hilfe digitaler SRL-Trainings zu verbessern, bedarf es weiterer Untersuchungen, z.B. zu digitalen, spielebasierten Lernumgebungen.

Digitalisierungsbezogene Kompetenzen von Lehrkräften: Erste Items der Skala „Digitalität und Bildung - Lehramt“ (dib-LA)

David Alexej Tobinski, Nils Nolte, Detlev Leutner
University of Duisburg-Essen, Germany

Theoretischer Hintergrund. Digitalisierungsbezogene Kompetenzen spielen in allen Lebensbereichen eine zunehmend große Rolle, und die damit verbundene Diskussion um eine digitale Bildung geht im

deutschsprachigen Raum bis hin zur Frage nach einer digitalen Grundbildung (Wolf & Koppel, 2017). Einigkeit herrscht in allen Diskussionen, dass es sich hinsichtlich einer digitalen Bildung (*digital literacy*) nicht um eine rein informatische Bildung handelt (u.a. Frankfurt Dreieck; Brinda et. al. 2019). Auf Europäischer Ebene besteht seit 2013 das Digital Competence Framework for Citizens (DigComp; Ferrari, 2013), welches sich in seiner jüngsten Version aus fünf Dimensionen zusammensetzt (DigComp 2.2; Vuorikari et. al., 2022). Auch die KMK-Strategie „Bildung in der digitalen Welt“ (KMK, 2016) orientiert sich für eine bundesweite Auslegung am Europäischen Rahmenkonzept. Für die Implementierung im Bildungssystem gilt es aktuell weiterhin, eine Operationalisierung der Konzepte der benannten Rahmenwerke in konkrete Kompetenzmodelle auf Länderebene zu übersetzen. Während dies in einigen Bundesländern auf der Ebene von Schülerinnen und Schülern bereits verbindlich erfolgt ist, z.B. der Medienkompetenzrahmen NRW (Medienberatung NRW, 2018), liegen auf der Ebene von Lehrerinnen und Lehrern national und international zunächst lediglich Orientierungsrahmen vor (Eickelmann, 2020; Redecker, 2017). Mit dem *integrativen Modell digitalisierungsbezogener Kompetenzen für die Lehramtsausbildung*, auch als UDE-Modell bezeichnet, haben Beißwenger et. al. (2020) eine Zusammenführung bestehender Rahmenwerke vorgenommen und für die erste Phase der Ausbildung des Lehramtes (LA) ausgearbeitet. Die Autoren identifizieren hierbei vier wesentliche Bereiche: (A) Fachspezifische, fächerübergreifende überfachliche Basiskompetenzen zu digitalen Technologien, (B.1) Fachspezifische, fächerübergreifende überfachliche Kompetenzen zum Unterrichten mit digitalen Technologien und zur Digitalisierung als Unterrichtsgegenstand, (B.2) Fachspezifische, fächerübergreifende überfachliche Kompetenzen für die Schul- und eigene Professionsentwicklung in der digitalisierten Welt, (C) kritisch-konstruktive Praxis und deren Reflexion.

Fragestellung. In einer ersten Phase der Operationalisierung des Modells interessieren wir uns für die Abbildbarkeit der Bereiche B.1 und B.2 des UDE-Modells, welche konform zu den Empfehlungen der KMK-Strategie angelegt werden kann. Dies bedeutet, dass folgende Kompetenzbereiche der KMK-Strategie eine Berücksichtigung finden: (1) Unterrichten, Erziehen, (3) Beurteilen und (4) Innovieren (KMK, 2016, 2019). Hierbei sind Unterrichten, Erziehen und Beurteilen dem Bereich B.1 im UDE-Modell zuzuordnen, Innovieren bezieht sich auf B.2 im UDE-Modell. Es stellt sich für uns die Frage nach der Formulierbarkeit eines Kompetenzniveaumodells.

Methode. Eine erste Pilotierungsstudie ist mit *244* Studierenden (*w* = 173/66/5) des Lehramtes der Universität Duisburg-Essen in der ersten Jahreshälfte 2022 durchgeführt worden. Der Altersdurchschnitt der Stichprobe lag bei 23.27 Jahren (*sd* = 4.85). Die Datenerhebung wurde innerhalb des Learning-Management-Systems *Moodle* (Version 3.7) als vollanonymisierter, computerbasierter Test realisiert. Die für diesen Test entwickelten 14 Items (12 Multiple-Choice/2 Single-Choice/4 Auswahloptionen) verteilen sich wie folgt auf die Kompetenzbereiche: B1./Unterrichten (m=5), B.1/Erziehen (m=3), B.1/Beurteilen (m=3) und B.2/Innovieren (m=3). Im Sinne einer *digital literacy* erfolgt die Datenanalyse der Testitems auf Basis der *Item-Response-Theory* (Rost, 1996). Für die Berechnung eines Partial-Credit-Models wurde das R-Paket eRm (Version 1.0-2) in der R Version 4.2.1 verwendet (Maier & Hatzinger, 2007). Für die Schätzung des

PC-Models mittels Conditional-Maximum-Likelihood-Methode sind die Daten zuvor um fehlende Werte bereinigt worden, was zu einem Gesamtsample von N=150 führte.

Ergebnisse und ihre Bedeutung. Alle 14 Items (*Cronbachs Alpha* = .61) sind zur Schätzung des Modells verwendet worden. Die Parameterschätzung konvergierte nach 35 Iterationen bei insgesamt 25 Parametern (Conditional log-likelihood: -1410.282). Die Infitschätzer liegen zwischen 0.80 und 1.14. und verteilen sich annähernd symmetrisch um den exakten Fit-Wert. Die Testitems streuen ausreichend auf dem Kontinuum des Personenparameters und es bilden sich angemessene Relationen in den Schwierigkeitsgraden ab. Hierbei zeigt sich, dass die Items zum Erziehen und Innovieren schon gut funktionieren, hinsichtlich der Schwierigkeit fallen insbesondere die Items zum Unterrichten ins Auge. Aktuell gilt es, den Item-Pool weiterhin zu erweitern und den Test für einen zukünftigen adaptiven Einsatz auszubauen.

Mobile Endgeräte im Unterricht – Welche Geräte sollten genutzt werden? Und wie?

Daniel Laumann, Maurice Krause, Fabienne Kremer, Barbara Leibrock, Malte Ubben, Cornelia Denz, Boris Forthmann, Gilbert Greefrath, Susanne Heinicke, Annette Marohn, Dörthe Masemann, Felix Reer, Thorsten Quandt, Elmar Souvignier, Stefan Heusler
Westfälische Wilhelms-Universität Münster, Germany

Theoretischer Hintergrund

Mobile Endgeräte gewinnen in den letzten Jahren zunehmend an Bedeutung für den Schulunterricht im Allgemeinen und die mathematisch-naturwissenschaftlichen Fächer im Speziellen (Liu et. al., 2017; Girwidz et al., 2019). Diese Entwicklung wurde durch die Einflüsse der COVID-19-Pandemie beschleunigt (Vargo et al., 2020). In den genannten Fächern werden mobile Endgeräte in vielfältiger Form für die Nutzung interaktiver Simulationen (Wieman et al., 2008), dynamischer Modellbildungssoftware (Arbain & Shukor, 2015) oder zur digitalen Messwerterfassung (Staacks et al., 2018) genutzt.

Insbesondere bei Smartphones werden sowohl schülereigene Geräte nach dem Prinzip *Bring Your Own Device*(BYOD) als auch schuleigene Geräte aus zentralen Gerätepools (POOL) genutzt. Ein wichtiger Vorteil des BYOD-Ansatzes liegt – auch aus ökonomischer und ökologischer Sicht – darin, dass (fast) alle Lernenden ohnehin über entsprechende Geräte verfügen (Bookhagen et al., 2020; Griffiths & Williams, 2018). Umgekehrt lässt sich mit dem Ziel eines fokussierten Lernens argumentieren, dass bei POOL-Geräten, z. B. soziale Medien nicht verfügbar sind, wodurch das Ablenkungspotential geringer ausgeprägt ist als beim BYOD-Ansatz (Al-Furaih & Al-Awidi, 2021). Ergänzend stellt sich die Frage, inwiefern – jenseits der Art der Geräte – konkrete Nutzungsregeln, wie eine zeitliche Begrenzung der Nutzung des Smartphones, die Lernprozesse beeinflussen.

Fragestellungen

Ziel unserer Untersuchung war eine empirische Analyse zum Vergleich von Lernprozessen und Effekten der beiden Ansätze BYOD und POOL. Die erste Forschungsfrage (FF1) adressiert dabei Unterschiede in den Lernprozessen bei den beiden Ansätzen im Hinblick auf Wissenserwerb, Interesse und kognitive Belastung. Forschungsfrage (FF2) untersucht anhand derselben Konstrukte, wie sich eine zeitlich begrenzte Nutzung von Smartphones nach dem BYOD-Ansatz auf Lernprozesse auswirkt.

Methode

Die Erhebung erfolgte im Rahmen von Workshops (ca. 5 Stunden) in den Fächern Chemie, Mathematik und Physik zum Thema „Elektromobilität“ unter lernbezogener Nutzung von Smartphones mit Klassen der Jahrgangsstufen 8 und 9. Die Klassen konnten an maximal drei Workshops teilnehmen. Jede der teilnehmenden Klassen wurde zufällig entsprechend der jeweiligen Bedingungen in zwei Gruppen aufgeteilt.

Die Untersuchung erfolgte in einem quasi-experimentellen Vergleichsgruppendesign (FF1 – Gruppe 1: BYOD-Ansatz, Gruppe 2: POOL-Ansatz; FF2 – Gruppe 1: BYOD zeitlich beschränkt, Gruppe 2: BYOD zeitlich unbeschränkt) mit N = 896 (FF1) bzw. N = 525 (aktueller Stand der Datenauswertung für FF2) Lernenden.

Die Analyse von Wirkungen der jeweiligen Ansätze der Smartphonenutzung wurde in einem Prä-Post-Design erhoben. Kriterien waren die fachspezifische Lernleistung (Wissenstests in den drei Fächern), das situationale Interesse unter Berücksichtigung gefühls- und wertbezogener Valenzen (Schiefele, 1999) sowie die kognitive Belastung der Lernenden (Paas, 1992). Zudem liegen für das situationale Interesse und die kognitive Belastung zusätzliche Längsschnittdaten (drei weitere Messzeitpunkte) über den Verlauf der Workshops vor.

Ergebnisse

Die vorliegenden Daten bezogen auf FF1 zeigen u. a. in der Entwicklung der fachlichen Lernleistung zwischen Prä- und Posttest keinen Gruppenunterschied zwischen BYOD- und POOL-Gruppe, z. B. Physik; einfaktorielle Varianzanalyse mit Messwiederholung (rmANOVA): $F(1, 340) = 0.38$, $p = .536$. Die vorliegenden Ergebnisse weisen für die schulische Praxis darauf hin, dass bei der Entscheidung zur Art der Smartphonenutzung vor allem organisatorische, ökonomische oder ökologische Aspekte leitend sein können.

Im Hinblick auf FF2 zeigen vorläufige Analysen, dass die unbeschränkte Nutzung der BYOD-Geräte im Unterricht keinen Nachteil gegenüber der beschränkten Nutzung für die Lernleistung besitzt. Zudem weisen die Analysen u. a. Vorteile der freien Nutzung hinsichtlich des situationalen Interesses der Lernenden auf, z. B. Physik; beschränkte Nutzung: $M = 2.50$, $SD = 0.57$, $n = 126$; unbeschränkte Nutzung: $M = 2.70$, $SD = 0.45$, $n = 127$; t-Test für unabhängige Stichproben, Mittelwerte über fünf Messzeitpunkte: $t(238.97) = -2.98$, $p = .003$, $d = 0.51$. Diese Erkenntnisse könnten in der konkreten Lehr-Lern-Situation zur Entlastung von Lehrkräften führen, da diese weniger Maßnahmen zur Einschränkung der Smartphonenutzung in die Unterrichtsplanung und -durchführung einbeziehen müssten.

Symposien

C-11: Individuelle und kontextuelle Einflussfaktoren der sozial-emotionalen Entwicklung in der Kindheit

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* S07 S00 D07

Individuelle und kontextuelle Einflussfaktoren der sozial-emotionalen Entwicklung in der Kindheit

Chair(s): **Elisa Oppermann** (Otto-Friedrich-Universität Bamberg, Germany)
*Diskutant*in(nen):* **Wilfried Smidt** (Leopold-Franzens-Universität Innsbruck)

Sozial-emotionale Kompetenzen gelten als zentrale Prädiktoren für eine erfolgreiche akademische sowie persönliche Entwicklung von Kindern und Jugendlichen (Denham et al., 2014; Moffitt et al., 2011). Zudem werden sozial-emotionale Kompetenzen aufgrund ihrer Schlüsselrolle im Umgang mit persönlichen sowie gesamtgesellschaftlichen Herausforderungen und Krisen, wie z.B. der COVID-19 Pandemie, auch als „Schlüsselkompetenzen des 21. Jahrhunderts“ bezeichnet (OECD Lernkompass 2030; Sylva et al., 2020). Der Förderung sozial-emotionaler Kompetenzen in der Kindheit kommt daher aktuell besondere Aufmerksamkeit zu. Aus theoretischer Perspektive ist die sozial-emotionale Entwicklung einerseits durch individuelle Eigenschaften des Kindes beeinflusst (z.B. das Temperament; vgl. Slagt et al., 2016) sowie andererseits abhängig von den Lernumgebungen des Kindes, insbesondere im Elternhaus und in Einrichtungen der frühen institutionellen Bildung und Erziehung (Bronfenbrenner & Morris, 2006). Auch im Kontext der Resilienzforschung werden individuelle sowie kontextuelle Einflussfaktoren unterschieden, die einen förderlichen oder hemmenden Einfluss auf die kindliche sozial-emotionale Entwicklung haben können (Fröhlich-Gildhoff & Rönnau-Böse, 2020; Prime et al., 2020). Bisher existieren jedoch kaum empirische Studien, die sowohl individuelle als auch kontextuelle Einflussfaktoren und deren Zusammenwirkung berücksichtigen. Das geplante Symposium widmet sich daher in drei Längsschnittstudien der Bedeutsamkeit individueller (z.B. Temperament) sowie kontextueller (z.B. Elternhaus und Kita) Einflussfaktoren und deren förderlichen bzw. hemmenden Einflüssen auf die sozial-emotionale Entwicklung von Kindern in der Altersgruppe von 0-9 Jahren.

Der erste Beitrag untersucht die Rolle des kindlichen Temperaments für den Umgang mit förderlichen und hemmenden Einflüssen im Elternhaus und deren Bedeutsamkeit für die kindliche sozial-emotionale Anpassung im Alter von 5 Jahren.

Der zweite Beitrag untersucht den Einfluss negativer Emotionalität (als ein Aspekt des kindlichen Temperaments) im Zusammenhang mit dem elterlichem Stresserleben auf das sozial-emotionale Verhalten von Kindern in der Einschulungsphase.

Der dritte Beitrag untersucht den Einfluss individueller Voraussetzungen sowie der kindlichen Lernumgebungen im Elternhaus und in Kindertageseinrichtungen für die Entwicklung sozial-emotionaler Probleme bei Kindern im Alter von 3-9 Jahren im Kontext der COVID-19 Pandemie.

Insgesamt tragen die Beiträge des geplanten Symposiums zu einem besseren Verständnis der differenziellen Bedeutsamkeit individueller und kontextueller Einflüsse auf die sozio-emotionale Entwicklung in der Kindheit bei. Zudem geben die Beiträge Einblick in förderliche sowie hemmende Einflüsse auf die sozial-emotionale Entwicklung in herausfordernden Umständen (z.B. Stress im

Elternhaus, Einschulung, COVID-19 Pandemie). Die Beiträge sowie deren Implikationen für Forschung und Praxis werden von Prof. Dr. Wilfried Smidt diskutiert.

Beiträge des Symposiums

Frühkindliches Temperament und sozio-emotionale Anpassung im Vorschulalter: Evidenz für Diathese-Stress oder differential susceptibility?

Jessica Willard[1], Anna Volodina[2], Manja Attig[1]
[1]Leibniz-Institut für Bildungsverläufe, Bamberg, [2]Otto-Friedrich-Universität Bamberg, Germany

Theoretischer Hintergrund. Frühes Temperament als Indikator für Sensitivität gegenüber Umweltbedingungen hat möglicherweise langfristige Konsequenzen für die sozio-emotionale Anpassung (Slagt et al., 2016). Ein „schwierigeres" frühes Temperament könnte einerseits vulnerabel für negative Umwelteinflüsse machen (Diathese-Stress-Ansatz). Alternativ könnte ein schwierigeres Temperament auch unter bestimmten Umweltbedingungen einen Vorteil darstellen: es könnte Ausdruck einer höheren Empfänglichkeit nicht nur für negative sondern auch für positive Umwelteinflüsse sein (differential susceptibility-Ansatz) (Belsky et al., 2007). Insbesondere für negative Emotionalität weist bisherige Forschung eher auf differential susceptibility hin (Slagt et al., 2016). Längsschnittstudien können aufzeigen, inwieweit dies auch über das Kleinkindalter hinaus gilt. Fragestellungen. Moderiert frühes Temperament die Effekte ausschließlich negativer Umwelteinflüsse (Diathese-Stress) oder auch positiver Umwelteinflüsse (differential susceptibility) auf sozio-emotionale Anpassung mit fünf Jahren? Methode. Für n = 1380 Säuglinge im mittleren Alter von sieben Monaten wurden in der Neugeborenenkohorte des Nationalen Bildungspanels (NEPS) aus Mutterinterviews negative Affektivität und regulatorische Kapazität erfasst anhand einer Kurzform des Infant Behavior Questionnaire–Revised (Bayer et al., 2015) und als Indikator für negative Umwelteinflüsse ein Familienrisikoscore gebildet (elterlicher Schulabbruch, Armutsrisiko, Vater arbeitslos, Eineltern-Haushalt, >3 Geschwister im Haushalt, mütterlicher Distress >1SD). Letzterer wurde aufgrund niedriger Prävalenz mehr als eines Risikos dichotomisiert. Ebenfalls zu diesem Messzeitpunkt wurde als ein zentraler positiver Umwelteinfluss mütterliche Sensitivität in einer standardisierten Spielsituation beobachtet (Linberg et al., 2019). Mit fünf Jahren wurden als Indikator für (geringe) sozio-emotionale Anpassung für diese Kinder vier SDQ-Subskalen von Müttern erhoben (Blossfeld & Roßbach, 2019). Fehlende Werte in Prädiktoren wurden mit kNN-Imputation behandelt (Kowarik & Templ, 2016). Ergebnisse und ihre Bedeutung. Vorläufige Regressionsmodelle kontrollieren für Alter in Monaten zum vorschulischen Messzeitpunkt, Geschlecht, Migrationshintergrund und Alter bei Eintritt in den Kindergarten. In einem ersten Set von Modellen wurden Haupteffekte negativen Affekts und regulatorischer Kapazität unter Berücksichtigung von mütterlicher Sensitivität und der Präsenz von Risiko untersucht. Hierbei zeigte sich, dass frühes Temperament ausschließlich ein Prädiktor der Skala „prosoziales Verhalten" war (ein positiver Zusammenhang für regulatorische Kapazität und ein negativer für negativen Affekt) und Umwelteinflüsse Prädiktoren von Verhaltensproblemen sowie Hyperaktivität waren. In weiteren Sets von Moderationsmodellen wurden einzeln Interaktionen zwischen den Temperamentsskalen und der Präsenz von Risiko bzw. mütterlicher Sensitivität getestet. Hierbei war nur eine einzige Interaktion signifikant (regulatorische Kapazität x mütterliche Sensitivität für die Skala „Probleme mit Gleichaltrigen"). Weiteres Proben dieser Interaktion zeigte, dass der moderierte Zusammenhang weder für den Diathese-Stress noch für den differential

susceptibility-Ansatz sprach. Insgesamt liefern diese vorläufigen Ergebnisse kaum Evidenz, dass frühe reaktive oder regulierende Temperamentsfacetten eine zentrale Stellung als Moderator von Umwelteinflüssen auf spätere vorschulische sozio-emotionale Anpassung haben. Dies kann u.a. mit der geringen Prävalenz von Risiken in der Stichprobe zusammenhängen. In Populationen, in denen eine größere Varianz an Risikofaktoren vorliegt, könnte eine moderierende Rolle frühen Temperaments offensichtlicher werden. Endgültige Analysen aggregieren Prädiktoren aus den ersten zwei Lebensjahren und berücksichtigen eine weitere Substichprobe von Familien mit fehlender Sensitivitätsbeobachtung aber höherer Risikobelastung.

Zusammenhänge zwischen elterlichem Stress, akademischen Problemen und sozial-emotionalen Kompetenzen bei Kindern mit hoher negativer Emotionalität in der Einschulungsphase

Catherine Gunzenhauser, Katja Berg
Universität Freiburg

Die Einschulung ist ein Einschnitt im Leben eines Kindes und seiner Eltern. Das Kind muss sich an neue Abläufe und Regeln gewöhnen, sich in einem neuen sozialen Umfeld bewähren und neuen Anforderungen gerecht werden (Blair & Raver, 2015). Studien finden häufig, dass vor der Einschulung gemessene akademische und sozial-emotionale Kompetenzen große Vorhersagekraft für die weitere Entwicklung im entsprechenden Bereich aufweisen (e.g., Pace et al., 2018). Gleichzeitig können Temperamentsmerkmale die schulische Anpassung beeinflussen. Insbesondere wurden für Kinder mit einer hohen Ausprägung des Temperamentsmerkmals der negativen Emotionalität größere Schwierigkeiten im akademischen Bereich (z. B. Nasvytienė, & Lazdauskas, 2021) sowie in der Beziehungsgestaltung zu Gleichaltrigen beschrieben (z. B. Hasenfratz et al., 2015).

Elterliche Unterstützung kann zur akademischen und sozial-emotionalen Entwicklung der Kinder beitragen; gleichzeitig kann erhöhtes elterliches Stresserleben die Unterstützung ihrer Kinder beeinträchtigen (vgl. McWayne et al., 2004). Es gibt Hinweise darauf, dass Kinder mit hoch ausgeprägter negativer Emotionalität besonders empfänglich für die Ausprägung elterlicher Unterstützung sind (cp. Shewark et al., 2021). Unserer Kenntnis nach wurde aber der Zusammenhang zwischen elterlichem Stress und der Entwicklung akademischer Probleme und sozial-emotionaler Kompetenz in der Einschulungsphase bisher nicht für Familien von Kindern mit hoch ausgeprägter negativer Emotionalität untersucht.

Die vorgestellte Studie untersucht die Zusammenhänge zwischen elterlichem Stresserleben in ihrer Elternrolle, schulischen Problemen der Kinder sowie Problemverhalten und prosozialem Verhalten der Kinder über die ersten acht Monate nach Schuleintritt hinweg.

Die Daten wurden im Rahmen eines größeren Projektes erhoben, in dem Forschungsfragen bezüglich eines Elterncoachings im Vordergrund standen. In die hier vorgestellten Analysen flossen Daten über N = 126 Erstklässler/innen (53% Mädchen, Alter zu Beginn der Studie: 6.48 Jahre, SD = 0.40 Jahre) sowie je ein Elternteil ein (97% Mütter). Ein Großteil der Eltern hatten für ihre Kinder höhere Werte im Temperamentsmerkmal „Negativer Affekt“ (Kurzfassung des Child Behavior Questionnaire; cp. Bayer et al., 2015; Putnam & Rothbart, 2006) als 85% einer Vergleichsstichprobe des Nationalen Bildungspanels (Blossfeld & von Maurice, 2011). Die Eltern berichteten in den ersten

acht Monaten nach Schuleintritt jeweils monatlich über die Intensität ihres Stresserlebens in der Elternrolle (cp. Crnic, 1991 & Booth, 1991), Einschätzung schulischer Probleme der Kinder (Drei Items über Probleme in den Bereichen Lesen, Schreiben und Mathematik, cp. Forschungsdatenzentrum am Institut zur Qualitätsentwicklung im Bildungswesen, 2010) sowie über Problemverhalten und prosoziales Verhalten der Kinder (Strenghts and Difficulties Questionnaire; Klasen et al., 2003).

Die Variablen enthielten zwischen 10% und 26% fehlende Werte auf Ebene der Skalen, bedingt größtenteils dadurch, dass Eltern zu späteren Zeitpunkten den Fragebogen nicht mehr ausfüllten.

Vorläufige Analysen mit listenweisem Fallausschluss ergaben zufriedenstellende interne Konsistenzen (Cronbach's Alpha zwischen .54 und .80). Multivariate lineare Regressionsanalysen zeigten, dass elterliches Stresserleben, schulische Probleme, Problemverhalten und prosoziales Verhalten der Kinder jeweils signifikant und positiv durch die Werte des Vormonats im jeweiligen Bereich vorhergesagt wurden. Dies galt auch unter statistischer Kontrolle von Geschlecht des Kindes, Teilnahme am Elterncoaching, und den Werten des Vormonats in den anderen untersuchten Bereichen. Lernprobleme wurden durch keinen der anderen Bereiche vorhergesagt. Andere bereichsübergreifende Zusammenhänge ergaben keine über die Zeit hinweg konsistenten Vorhersagemuster.

Die vorläufigen Ergebnisse weisen darauf hin, dass bei Kindern mit hoch ausgeprägter negativer Emotionalität bereits zu Beginn der Schulzeit ausgeprägte Muster an Lernproblemen, Problemverhalten, prosozialem Verhalten und elterlichem Stresserleben zentral für die weitere Entwicklung in diesen Bereichen sind und unterstützen damit die Bedeutsamkeit vorschulischer Förderung.

Für die Konferenz planen wir, Ergebnisse von Cross-lagged-Panelanalysen unter Verwendung von Full Information Maximum Likelihood zu präsentieren. Die Bedeutung der Elternperspektive als Informationsquelle für alle erhobenen Studienvariablen wird aus methodischer und theoretischer Perspektive diskutiert.

Sozial-emotionale Entwicklungsprobleme von Kindern in der COVID-19 Pandemie: Die Rolle häuslicher und institutioneller Lernumgebungen

Elisa Oppermann[1], Sabine Blaurock[1], Lysann Zander[2], Yvonne Anders[1]
[1]Otto-Friedrich-Universität Bamberg, Germany, [2]Leibniz Universität Hannover

Die COVID-19 Pandemie sowie die Einschränkungen zur Eindämmung des Virus traf Kinder im Vorschulalter hart: Durch wiederholte Lockdowns war vielen Kindern der Zugang zu institutioneller früher Bildung und Erziehung in Kindertageseinrichtungen (Kitas) zeitweise verwehrt, soziale Kontakte zu Peers waren eingeschränkt und auch das Familienleben wurde durch die Maßnahmen beeinträchtigt (Feinberg et al., 2022; Langmeyer et al., 2020; Oppermann et al., 2021). Vor dem Hintergrund empirischer Befunde zur Bedeutsamkeit der Beziehungen innerhalb sowie außerhalb der Familie für die Entwicklung von Fähigkeiten im Umgang mit den eigenen Emotionen sowie im Umgang mit anderen (Ahnert et al., 2006), können sozial-emotionale Entwicklungsprobleme von Kindern als Folge der Pandemie angenommen werden. Studien legen nahe, dass derartige Krisensituationen jedoch nicht bei allen Kindern zu Entwicklungsproblemen führen, sondern negative

Folgen von individuellen sowie kontextuellen Resilienzfaktoren abhängig sind (Prime et al., 2020). Resilienzfaktoren auf individueller Ebene umfassen Voraussetzungen des Kindes sowie Ressourcen der Familie, z.B. der sozio-ökonomische Status (Anders et al., 2022). Auf kontextueller Ebene wird in der Forschung insbesondere die Bedeutung der Qualität der häuslichen sowie institutionellen Lernumgebungen diskutiert (Anders, 2013; Feinberg et al., 2022; Sylva et al., 2020).

Die vorliegende Studie untersucht die Bedeutung individueller Merkmale sowie der Qualität der häuslichen sowie institutionellen Lernumgebungen als kontextuelle Merkmale für die sozial-emotionale Entwicklung von Kindern während der COVID-19 Pandemie.

Die Studie basiert auf Daten von 228 Familien, die im Rahmen des Bundesprogramms „Sprach-Kitas“ zu Messzeitpunkt (MZP) 1 im Herbst 2019 sowie erneut zu MZP2 im Winter 2021/2022 befragt wurden. Die Kinder waren zu MZP1 im Schnitt 5.13 Jahre und zu MZP2 im Schnitt 6.99 Jahre alt. Zu beiden MZP schätzten Eltern die emotionalen Probleme (5 Items; McDonald’s Omega (ω) MZP1=.56; ωMZP2=.72) sowie Probleme in Peer-Beziehungen (kurz: Peer-Probleme; 5 Items; ωMZP1=.58; ωMZP2=.69) ihrer Kinder ein. Als individuelle und familiale Einflussfaktoren wurden das Alter und Geschlecht der Kinder, die Familiensprache sowie das Bildungsniveau und das Nettoäquivalenzeinkommen der Eltern berücksichtigt. Die Qualität der häuslichen Lernumgebung wurde über das Familienklima (13 Items; ωMZP2=.86) sowie die Häufigkeit früher Bildungsaktivitäten im Elternhaus (6 Items; ωMZP2=.73) operationalisiert. Beobachtungen zur Einschätzung der pädagogischen Qualität mit der SSTEW (Anders et al., 2017) liegen zu MZP1 (2019) für die von den Kindern besuchten Kitas vor. Strukturelle Rahmenbedingungen der Kitas wurden in den Analysen kontrolliert.

Die Ergebnisse auf individueller Ebene zeigten eine signifikante Zunahme emotionaler Probleme von MZP1 (M=1.55) zu MZP2 (M=2.19; t(222) = 4.6; p<.001). Auch Peer-Probleme nahmen im Mittel zu, aber der Unterschied war statistisch nicht signifikant (MMZP1=1.08; MMZP2=1.19, p>.05). Regressionsanalysen zeigten, dass Jungen – im Vergleich zu Mädchen – sowie Kinder aus Familien mit geringerem Bildungsniveau – im Vergleich zu Kindern aus Familien mit höherem Bildungsniveau – problematischere Peerbeziehungen entwickelten. Die Entwicklung emotionaler Probleme hingegen wurde nicht durch individuelle Merkmale vorhergesagt. Mit Blick auf die familiale Lernumgebung zeigte sich weder ein Einfluss der Häufigkeit elterlicher Bildungsaktivitäten noch des Familienklimas auf die Entwicklung emotionaler und Peer-Probleme der Kinder von MZP1 zu MZP2. Es zeigten sich hingegen Effekte der Kita-Qualität auf die Entwicklung von Peer Problemen: Je höher die von den Kindern erfahrene pädagogische Qualität in der Einrichtung zu MZP1, desto geringer fiel die Entwicklung von Peer-Problemen bei den Kindern zwischen MZP1 und MZP2 aus (ß=-.29*). Für die emotionalen Probleme zeigte sich kein signifikanter Effekt der Kita-Qualität (ß=-.07, p>.05).

Die Ergebnisse des vorliegenden Beitrags deuten darauf hin, dass sozial-emotionale Probleme im Zuge der COVID-19 Pandemie nicht bei allen Kindern gleichermaßen zunahmen, sondern insbesondere bei Jungen und in Familien mit geringerem Bildungsniveau (siehe auch Ravens-Sieberer et al., 2021). Eine höhere frühpädagogische Qualität in Kitas hatte einen protektiven

Einfluss. Die Befunde unterstreichen die Bedeutsamkeit der Kita-Qualität für die kindliche sozial-emotionale Entwicklung – insbesondere in Krisen wie der COVID-19 Pandemie.

C-12: Digitale Transformation von Schulen der Sekundarstufe II in Deutschland und der Schweiz aus Schulentwicklungsperspektive

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* T03 R02 D26

Digitale Transformation von Schulen der Sekundarstufe II in Deutschland und der Schweiz aus Schulentwicklungsperspektive

Chair(s): **Stephan Schumann** (Universität Konstanz, Germany)
*Diskutant*in(nen):* **Andreas Lachner** (Universität Tübingen)

Theoretischer Rahmen

Spätestens mit der Veröffentlichung der Ergebnisse der ICIL-Studie 2013 (Bos et al., 2014) rückte die digitale Transformation im Bildungswesen in den Fokus. Im Zuge der COVID19-Pandemie und der damit verbundenen digitalisierungsbezogenen Herausforderungen gewann das Thema dann nochmals an Gewicht (Huber et al., 2020; Feldhoff et al., 2022). Wie vielschichtig und umfassend dieser Wandel für Bildungsorganisationen ist, zeigen einschlägige Modelle und Befunde (z.B. Ilomäki & Lakkala, 2018; Petko et al., 2018; Eickelmann & Gerick, 2018). Bei der erfolgreichen Gestaltung dieses Schulwandels und der effektiven Umsetzung entsprechender Maßnahmen nimmt die Schulleitung eine zentrale Rolle ein (Gerick & Eickelmann, 2019).

Übergeordnete Fragestellung

Das Symposium setzt sich mit der übergeordneten Fragestellung auseinander, wie die digitale Transformation in deutschen und schweizerischen Schulen der Sekundarstufe II wahrgenommen wird und insbesondere, welchen Beitrag Schulleitungen leisten können, damit dieser Schulentwicklungsprozess erfolgreich gestaltet werden kann.

Beiträge des Symposiums

Der erste Beitrag geht der Frage nach, inwieweit Schulleitungen auf Digitalisierungsprozesse Einfluss nehmen können und welche Maßnahmen in diesem Zusammenhang von besonderer Bedeutung sind. Im Zuge eines Online-Surveys wurden hierzu n = 3.872 unterschiedliche Akteur:innen berufsbildender Schulen in Baden-Württemberg befragt. Der inhaltliche Fokus liegt dabei auf Distance Learning - Prozessen, deren besondere Bedeutung insbesondere im Zuge der COVID19-Pandemie zum Vorschein kam.

Der Frage, ob ein transformationaler Führungsstil als Prädiktor für eine digitale Schulkultur im Sinne formeller und informeller Kollaboration, Zielklarheit und der Bedeutung digitaler Medien fungiert, geht der zweite Beitrag des Symposiums nach. Als Datenbasis dient eine breit angelegte Befragung in der gesamten Schweiz, an der insgesamt n = 2.265 Lehrpersonen aus n = 112 Schulen teilgenommen haben.

Der dritte Beitrag zielt auf die mehrdimensionale Betrachtung der digitalen Transformation beruflicher Schulen und vergleicht dabei die Perspektiven von Schulleitungen in der Schweiz und in Deutschland. Es wird insbesondere untersucht, welchen Beitrag die Dimension „Strategie und Führung“ zur Wahrnehmung des digitalisierungsbezogenen Entwicklungsstands leistet. Zu diesem Zweck werden Daten von n = 466 Schulleitungsmitglieder aus der Schweiz sowie n = 117 Schulleiter:innen aus Baden-Württemberg verwendet.

Die digitale Transformation von Schulen der Sekundarstufe II wird im Symposium somit aus verschiedenen Blickwinkeln betrachtet, wodurch ihr umfassender Charakter unterstrichen wird. Aktuelle Entwicklungen wie insbesondere der Umgang mit pandemiebedingten Herausforderungen werden dabei ebenfalls berücksichtigt.

Beiträge des Symposiums

Digital Distance Learning als Herausforderung und Chance für Schulleitungen

Jan Delcker, Dirk Ifenthaler
Universität Mannheim

Theoretische Fundierung

Die Auswirkungen weitreichender Digitalisierungsprozesse stellen eine der zentralen Herausforderungen für die Schulleitungen an berufsbildenden Schulen dar. Die Potentiale, welche die Digitalisierung für die Bildungseinrichtung „Schule“ birgt, traten in den letzten Jahren vor allem unter dem Einfluss der COVID-19 Pandemie in den Vordergrund (Schneider et al., 2021). Gleichsam wurde deutlich, welche Voraussetzungen erfüllt sein müssen, um berufsbildende Schulen durch digitalisierte administrative Prozesse sowie digitale Lehr- und Lernkonzepte weiterzuentwickeln (Delcker & Ifenthaler, 2021). Neue Aufgabenbereiche und Kompetenzen für die Schulleitungen werden nicht nur in der Praxis ersichtlich (Lindqvist & Pettersson, 2019), sondern sie lassen sich auch aus denjenigen theoretischen Modellen herleiten, in denen die Digitalisierung als wich-tiges Element zukunftsfähiger Schulentwicklung aufgenommen wird (Ilomäki & Lakkala, 2018; Labusch et al., 2020). Die Entscheidungen von Schulleiter:innen haben Einfluss auf die Zufrie-denheit und die erbrachte Leistung von Lehrpersonen (Mulford, 2003). Sie sind verantwortlich dafür, dass Lehrpersonen die Möglichkeiten erhalten, ihre Kompetenzen auszubauen (Huang et al., 2020). Aus den Strukturen der Bildungseinrichtung entsteht ein entscheidender Einflussfaktor für den Lernerfolg von Schüler:innen (Wang & Degol, 2016).

Fragestellung

Mit dem Projekt Check-up Distance Learning sollte unter anderem den Fragen nachgegangen werden, auf welche Weise Schulleitungen mittels datenbasierter Instrumente auf Digitalisierungsprozesse an ihren Schulen Einfluss nehmen und welche Maßnahmen besonders wichtig sind, um die Digitalisierung als Teil von Schulentwicklung effektiv zu implementieren.

Methode

Mit Hilfe eines digitalen Befragungsinstruments wurden die Daten von N = 3.872 Interessenvertreter:innen aus dem Umfeld der Berufsbildenden Schulen in Baden-Württemberg erhoben. Die befragten Lehrpersonen, Schüler:innen, Schulleitungen, Ausbildenden und Erziehungsberechtig-ten konnten ihre Einschätzung zu acht Bereichen der Schulentwicklung abgeben (Balzer & Schorn, 2021). Die Datenerhebung fand unter den Eindrücken der COVID-19 Pandemie und den daraus resultierenden Distance Learning Abläufen statt. Neben der Analyse der quantitativen Daten wurden auch N = 1.493 offene Antworten der Proband:innen ausgewertet.

Ergebnisse und ihre Bedeutung

Insbesondere bei der Einführung von Distance Learning Prozessen votiert ein hoher Anteil der Beteiligten (429 Aussagen) für klare Regeln und Vorgaben seitens der Schulleitungen. Diese Vorgaben bilden ein entscheidendes Grundgerüst, welches Kommunikationswege und Zuständigkei-ten abbildet. Bezeichnend für den Wunsch nach klaren Vorgaben ist der Kontrast zu den Freiheiten, die Lehrpersonen im Rahmen der Unterrichtsgestaltung im Allgemeinen zur Verfügung ste-hen (Kerres, 2020). Eine wichtige Aufgabe der Schulleitungen ist es, dieses Spannungsfeld aus-zutarieren, um durch Vorgaben Rahmenbedingungen zu schaffen, welche auf der Ebene des Un-terrichts individuelle Handlungsformen zulassen. Hierbei ist zu bedenken, dass die Schulleitun-gen selbst den lokalen Gegebenheiten und formalen Vorgaben der Schulträger Rechnung tragen müssen. Dieser Zusammenhang schlägt sich auch in der Forderung nach besserer IT-Infrastruktur nieder. Neben der Notwendigkeit ausreichender Kapazitäten zur Datenübertragung fehlt es an den Berufsbildenden Schulen an einer nachhaltigen Lösung, um Lehrpersonen und Schüler:innen mit den geeigneten Endgeräten sowie Softwarelösungen auszustatten. Im Vordergrund steht der Wunsch der Schüler:innen, über Live-Video-Formate unterrichtet zu werden. Über schulinterne Strategien, die in partizipativen Prozessen, auch unter Einbezug der ausbilden-den Betriebe, ausgearbeitet werden können, kann sichergestellt werden, dass Lehrpersonen dieser Anforderung beim Distance Learning nachkommen und Schüler:innen diese Angebote auch wahrnehmen. Entfremdungstendenzen zwischen den verschiedenen Akteur:innen, welche im Diskurs um die Digitalisierung hervorgebracht werden (Arnold, 2020), kann durch eine fest in der Schule verankerte Feedbackkultur zusätzlich begegnet werden. Die Schulleitungen nehmen im Austausch zwischen den Akteur:innen innerhalb der Schule eine entscheidende Rolle ein. Im Hinblick auf die Digitalisierung bedeutet dies, nicht nur Personen zusammen zu bringen, sondern auch schulische System zu vernetzen, um Daten aus multiplen Prozessen nutzbar zu machen. Im Sinne vernetzter Schulentwicklung lässt sich dieses Schulnetz über die Einzelschule hinaus erweitern, um verschiedene Schulstandorte, ausbildende Betriebe und zusätzliche Kooperationspartner in Schulentwicklungsprozesse einzubinden.

Transformationale Führung und digitale Schulkultur: Empirische Befunde einer repräsentativen Bestandsaufnahme in der Schweiz

Maria-Luisa Schmitz, Dominik Petko
Universität Zürich

Theoretische Fundierung

In der Forschung konnte wiederholt gezeigt werden, dass Führungsstile, die Bottom-Up und Top-Down Prozesse vereinen, besonders effektiv für die Technologieintegration in Schulen sind (Pet-ko et al., 2015). Ein Führungsstil, der diese auf den ersten Blick gegensätzlichen Ansätze verbindet, ist der der transformationalen Führung. Er umfasst sowohl Bottom-Up-Komponenten wie intellektuelle Stimulierung und individuelle Unterstützung von Mitarbeitenden (Daniëls et al., 2019) als auch Top-Down-Aspekte wie idealisierter Einfluss und inspirierende Motivierung. Verschiedene Studien verweisen auf positive Effekte transformationaler Führung auf diverse Indikatoren der Technologieintegration (Chen, 2013; Ottestad, 2013; Seyal, 2015; Vermeulen et al., 2015; Vermeulen et al., 2017; Yamamoto & Yamaguchi, 2019). Die meisten dieser Studien bleiben jedoch bei der Analyse der direkten Effekte von transformationaler Führung und Tech-nologieintegration stehen, ohne weitere mediierende Variablen einzubeziehen. Dass hier weitere Einflüsse anzunehmen wären,

zeigt z.B. die Studie von Petko et al. (2018), in der herausgefunden wurde, dass Leitungshandeln im Zusammenhang mit anderen Aspekten einer digitalen Schulkul-tur gesehen werden muss, wie etwa formeller und informeller Kollaboration der Lehrpersonen, Zielklarheit in Bezug auf digitale Medien und die Wichtigkeit digitaler Medien in der Schulstrate-gie. Daneben existieren auch qualitative Studien, die zeigen können, dass ein transformationaler Führungsstil der Schulleitung einen positiven Einfluss auf verschiedene Aspekte der digitalen Schulkultur haben könnte (Laschou, 2018; Ruloff & Petko, 2021). Folglich wäre es lohnenswert diesen Zusammenhang auch quantitativ zu untersuchen.

Hypothesen

Basierend auf den bestehenden Befunden können folgende Hypothesen aufgestellt werden:

1. Transformationale Führung korreliert signifikant positiv mit den verschiedenen Dimensio-nen der digitalen Schulkultur (formelle und informelle Kollaboration, Zielklarheit und Wich-tigkeit digitaler Medien).

2. Transformationale Führung ist ein signifikanter und positiver Prädiktor der verschiedenen Dimensionen einer digitalen Schulkultur (formelle und informelle Kollaboration, Zielklarheit und Wichtigkeit digitaler Medien).

Methode

Im Rahmen einer nationalen Fragebogenstudie in der Schweiz wurden alle Lehrpersonen der Sekundarstufe II, die in der vorletzten Jahrgangsstufe unterrichten, eingeladen an einer Online-Umfrage teilzunehmen.

In dieser Umfrage sollten die Lehrkräfte auf einer fünfstufigen Skala (1 = „Nie“ bis 5 = „Immer“) angeben, wie häufig ihre Schulleitung typisches Verhalten eines transformationalen Führungsstils zeigt (7 Items; Skala von Carless et al., 2000). Weiterhin wurden sie gebeten, die formelle und informelle Kollaboration, die Zielklarheit und die Wichtigkeit digitaler Medien an ihrer Schule auf einer fünfstufigen Skala (1 = „Stimme gar nicht“ bis 5 = „Stimmt völlig“; Skalen von Petko et al., 2018 mit je 3 Items) einzuschätzen.

Die Stichprobe umfasst 2.265 Lehrkräfte aus 112 Schulen. Die Hypothesen wurden mithilfe von Multilevel-Korrelationen und einem Within-only Multilevel Strukturgleichungsmodell geprüft. Dabei wurde transformationale Führung als within-Prädiktor für die vier genannten Dimensionen der digitalen Schulkultur geprüft.

Ergebnisse und ihre Bedeutung

Die Multilevel-Korrelationen zeigen, dass transformationale Führung signifikant und positiv mit formeller Kollaboration ($r = .36$, $p < .001$), informeller Kollaboration ($r = .33$, $p < .001$), Zielklar-heit ($r = .46$, $p < .001$) und Wichtigkeit digitaler Medien ($r = .38$, $p < .001$) zusammenhängt. Eben-so ist transformationale Führung in dem Strukturgleichungsmodell ein signifikant positiver Prädiktor für alle vier Dimensionen der digitalen Schulkultur. Die Ergebnisse der vorliegenden Stu-die stützen die Annahme, dass mit einem transformationalen Führungsstil Schulleitungen eine Schulkultur mit einer funktionierenden Kollaboration unter den Lehrkräften in Bezug auf digitale Medien schaffen, wo die Ziele in Bezug auf die digitale Transformation klar sind und digitale Medien einen hohen Stellenwert

haben. Diese Befunde stehen im Einklang mit vorherigen qualitativen Studien, die zeigen konnten, dass Schulleitende mit einem transformationalen Führungs-stil eine positivere Sicht auf schulische Innovationen mit digitalen Medien haben, klarere Ziele für ihre Schule in Bezug auf digitale Medien formulieren und den digitalen Wandel an ihren Schulen vorantreiben (Laschou, 2018; Ruloff & Petko, 2021).

Digitale Transformation beruflicher Schulen aus Sicht von Schulleitungen in der Schweiz und in Deutschland – Welche Rolle spielen Strategie und Führung?

Andreas Harder[1], Stephan Schumann[1], Serge Imboden[2]
[1]Universität Konstanz, [2]HES-SO Wallis

Theoretische Fundierung

Spätestens mit der Veröffentlichung der ICIL Studie 2013 wurden sowohl in der Schweiz als auch in Deutschland digitale Entwicklungspotenziale sichtbar (u.a. Nutzungshäufigkeit von ICT im Unterricht) (Bos et al., 2014; Eickelmann et al., 2014). Dies führte zu bildungspolitischen Maßnahmen, wie beispielsweise der Veröffentlichung des Lehrplans 21 in der Deutschschweiz sowie der Strategie „Bildung in der digitalen Welt“ (KMK, 2016) und der Verabschiedung des DigitalPakts Schule in Deutschland. Durch den „exogenen Schock“ der COVID19-Pandemie wurden entsprechende Entwicklungsbedarfe nochmals verdeutlicht und das Thema der digitalen Transformation generell nochmals in den Fokus gerückt (SWK, 2021; Harder et al., in Druck).

Verschiedene theoretische Ansätze zeigen, wie vielschichtig die digitale Transformation für Schu-len ist (u.a. Eickelmann & Gerick, 2018; Ilomäki & Lakkala, 2018). Das unserer Studie zugrunde gelegte Maturity Model for Educational Organizations grenzt sechs inhaltliche Dimensionen von-einander ab: Strategie und Führung, Ausstattung und Technik, Organisation, Mitarbeitende, Kultur sowie Digitales Lehren und Lernen (Egloffstein et al., 2019). Bei der erfolgreichen Gestaltung dieses Wandels nimmt die Schulleitung eine zentrale Rolle ein (Gerick & Eickelmann, 2019).

Fragestellung

Der Beitrag geht drei Forschungsfragen nach: Zum einen soll untersucht werden, inwieweit Schul-leitungen aus der Schweiz und Deutschland die verschiedenen inhaltlichen Dimensionen der digi-talen Transformation wahrnehmen (1) und ob diesbezüglich Unterschiede zwischen den beiden Ländern festgestellt werden können (2). Zum anderen stellt sich die Frage, welchen Erklärungsbeitrag die Dimension „Strategie und Führung“ im Sinne des Schulleitungshandelns für die Wahrnehmung des digitalisierungsbezogenen Entwicklungsstands der Schule leistet (3).

Methode

Die Daten zur Beantwortung der Fragestellungen entstammen zwei Online-Erhebungen: An den Befragungen haben n = 466 Personen aus der Schweiz (Erhebungszeitraum: April/Mai 2021) resp. n = 117 Schulleiter:innen aus Baden-Württemberg (Erhebungszeitraum: Februar 2022) teil-genommen. Im Gegensatz zur Stichprobe aus Deutschland umfasst die schweizerische Stichprobe auch Mitglieder der erweiterten Schulleitung. Als Analyseverfahren werden u.a. lineare Regressi-onsmodelle mit Standardfehlerkorrektur verwendet.

Ergebnisse und ihre Bedeutung

Die digitale Transformation wird sowohl gesamthaft als auch in ihren Teildimensionen sowohl in der Schweiz als auch in Deutschland (eher) positiv wahrgenommen; die Mittelwerte liegen durchgängig über der theoretischen Mitte. Im Hinblick auf die Priorisierung der digitalen Transformation ($p = .002$; $d = 0.31$) und die positiven Einstellungen der Lehrpersonen gegenüber ICT ($p < .001$; $d = 0.42$) liegt die Wahrnehmung in Deutschland signifikant über jener in der Schweiz. Insbesondere beim technischen ($p < .001$; $d = -0.65$), aber auch hinsichtlich des pädagogischen IT-Supports ($p = .004$; $d = -0.29$) zeigt sich dagegen ein umgekehrtes Bild. Multivariat wird ersichtlich, dass die Dimension „Strategie und Führung“ in beiden Ländern – neben anderen Prä-diktoren – inkrementell zur Varianzerklärung des wahrgenommenen digitalisierungsbezogenen Status Quo der Schulen beiträgt.

Für die zum Teil differenzierte Wahrnehmung der inhaltlichen Dimensionen der digitalen Transformation in beiden Ländern sind verschiedene Gründe denkbar, wie zum Beispiel unterschiedliche organisatorische Regelungen hinsichtlich der Zuständigkeit des IT-Supports. Im Einklang mit anderen Studienergebnissen (z.B. Gräsel et al., 2020; Dexter, 2018) unterstreichen die Befunde darüber hinaus die Relevanz der Schulleitung für die erfolgreiche Gestaltung der digitalen Transformation. Um zukünftige Herausforderungen bewältigen und Potenziale im Sinne dieses Wan-dels nutzen zu können, bedarf es einer nachhaltigen Schulentwicklung. Es ist eine zentrale Auf-gabe der Schulleitung, diese Prozesse aktiv zu begleiten und förderliche Rahmenbedingungen zu gewährleisten (vgl. u.a. Gerick et al., 2019; SWK, 2022). Vor diesem Hintergrund soll im Rah-men dieses Beitrags auch diskutiert werden, in welchem Ausmaß Schulleitungen (teil-)autonom auf Einzelschulebene agieren können bzw. sollten.

C-13: Evidence-based teacher education – Effekte von evidenzbasierten Lehr-Lern-Prinzipien auf die Wirksamkeit videobasierter Lernumgebungen in der Lehrkräftebildung

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* R11 T00 D01

Evidence-based teacher education – Effekte von evidenzbasierten Lehr-Lern-Prinzipien auf die Wirksamkeit videobasierter Lernumgebungen in der Lehrkräftebildung

Chair(s): **Bernadette Gold** (Technische Universität Dortmund, Germany), **Andreas Gegenfurtner** (Universität Augsburg)
*Diskutant*in(nen):* **Inga Glogger-Frey** (Universität Erfurt)

Zahlreiche Studien belegen einen positiven Effekt von Unterrichtsvideos auf professionelle Kompetenzen von (angehenden) Lehrkräften (Gaudin & Chaliès, 2015; Steffensky & Kleinknecht, 2016). Dabei ist eine adressat:innenorientierte sowie instruktional gut begründete Gestaltung videobasierter Lehr-Lern-Umgebungen für den Kompetenzerwerb zentral. Gleichzeitig ist erstaunlich, dass vergleichsweise wenige Studien (z. B. Moreno & Valdez, 2007; Kumschick et al., 2017; Seidel et al., 2013) explizit auf Theorien und Empirie zum Lernen mit Multimedia (z. B. Cognitive Load Theorie, Theorie zum multimedialen Lernen) zurückgreifen, um konkrete Analyseaufgaben, Begleitmaterial oder Unterstützungsmöglichkeiten zu generieren und schließlich zu evaluieren.

Im Sinne des Tagungsthemas „Bildung zwischen Unsicherheit und Evidenz“ sollte auch eine evidenzbasierte Lehrkräftebildung auf gesicherte Evidenz und Theorien zurückgreifen, um videobasierte Lehr-Lern-Umgebungen zu gestalten. Dementsprechend werden in dem Symposium Arbeiten vorgestellt, die aktuelle und gut gesicherte Evidenz aus der Lehr-Lern-Forschung sowie Multimedia-Forschung auf den instruktionalen Einsatz von Unterrichtsvideos übertragen und somit eine evidenzorientierte Lehrkräftebildung vorantreiben. Alle Beiträge untersuchen Konzepte zur videobasierten Förderung der professionellen Wahrnehmung oder des professionellen Wissens von angehenden Lehrkräften, weisen jedoch hinsichtlich der Forschungsdesigns (Experimente und Kurztrainings, längere Interventionen) und der Messinstrumente (videobasierte Testverfahren, inhaltsanalytische Kodierungen schriftlicher Videoanalysen, Klick-Aufzeichnungen während der Videobetrachtung mit anschließendem Interview) eine gewisse Vielfalt auf.

1. Im ersten Beitrag werden die Möglichkeiten für die Förderung der professionellen Wahrnehmung anhand eines *Emphasis-Shift-Trainings* kombiniert mit Erkenntnissen aus der Forschung zum beispielbasierten Lernen untersucht.
2. Der zweite Beitrag bezieht sich auf die Cognitive Load Theorie und untersucht in zwei experimentellen Studien Möglichkeiten zur Reduktion des cognitive load beim videobasierten Lernen von nonverbalen Klassenführungsstrategien – Signaling, Segmenting und das Ausblenden des Videotons.
3. Im dritten Beitrag wird auf Grundlage der Forschung zum Lernen aus Lösungsbeispielen in einer Interventionsstudie untersucht, ob schriftliche Beispielanalysen von Expert*innen Studierende beim Erwerb professionellen Wissens und professioneller Wahrnehmung unterstützen.
4. Aufbauend auf aktuellen Studien zu Effekten von Prompts und unterschiedlicher Beobachtungsaufgaben untersucht der vierte Beitrag, ob sich eine konkrete Aufgabenstellung vor einer Videoanalyse oder Prompts während der Videoanalyse als wirksamer für die Förderung der professionellen Wahrnehmung erweisen.

Die Ergebnisse legen die Schlussfolgerung nahe, dass der Transfer auf den hochkomplexen Gegenstand des Unterrichtens nicht uneingeschränkt möglich ist und dass weitere Studien notwendig sind, um evidenzbasierte Gelingensbedingungen für den Ziel- und adressat:innenangepassten Einsatz von Unterrichtsvideos zu identifizieren, wofür das Symposium Impulse geben möchte.

Beiträge des Symposiums

Emphasis-Shift-Training und Selbsterklärungsprompts als Unterstützungsmaßnahmen in einem Professional-Vision-Kurztraining

Monika Martin[1], Meg Farrell[2], Werner Rieß[3], Tina Seidel[2], Alexander Renkl[1]
[1]Albert-Ludwigs-Universität Freiburg, [2]Technische Universität München, [3]Pädagogische Hochschule Freiburg

Wenn Unterrichtsvideos zur Förderung der Professional Vision (Seidel & Stürmer, 2014) von Lehrkräften eingesetzt werden, ist die Einbettung der Videos in eine geeignete Lernumgebung entscheidend für den Lernerfolg (Seidel et al., 2013). Vor allem unerfahrenen Lehrkräften fällt es ohne zusätzliche Unterstützung oft schwer, in der Menge an unmittelbar hintereinander auftretenden Ereignissen relevante Ereignisse und deren Folgen für den Lernprozess der Schülerinnen und Schüler zu erkennen (Star & Strickland, 2008).

Für andere Berufe, in denen die Auszubildenden lernen müssen, flexibel auf ungewohnte, komplexe Situationen zu reagieren und gleichzeitig viele einströmende Informationen zu verarbeiten (z.B. Piloten), wurde das Emphasis-Shift-Training (EST) entwickelt (Gopher et al., 1989). Dabei bleibt die auszuführende Tätigkeit in ihrer Komplexität erhalten (z.B. in einem Simulator), die Lernenden fokussieren jedoch in mehreren Durchgängen jeweils einen anderen Aspekt.

Das Lernen mit Beispielen (wie z.B. Unterrichtsvideos) kann außerdem durch Selbsterklärungsprompts angereichert werden (Renkl & Eitel, 2019). Durch die Prompts werden Lernende dazu aufgefordert, generelles Wissen (z.B. über zugrundeliegende Prinzipien) auf eine konkrete Beispielsituation zu beziehen. Dabei haben sich fokussierte Prompts, die bereits eine Struktur bieten, als besonders effektiv erwiesen (Bertold et al., 2009).

In der vorliegenden Studie wurde untersucht, ob ein Emphasis-Shift-Training Lehramtsstudierende dabei unterstützen kann, relevante Ereignisse in einem Unterrichtsvideo zu erkennen und zu interpretieren. Außerdem sollte überprüft werden, ob fokussierte Selbsterklärungsprompts einen höheren Lernerfolg erzielen als offene Prompts.

Methode

An der Untersuchung nahmen 134 Biologie-Lehramtsstudierende (MAlter = 22.0, SD = 2.85) im Rahmen eines Seminars teil. Die Einheit bestand aus drei Teilen, die die Studierenden jeweils im Abstand von einer Woche online zuhause bearbeiteten: Das Emphasis-Shift-Training und die Art der Selbsterklärungsprompts wurden in einem einstündigen Kurztraining zum Thema Kleingruppentutoring variiert. Jeweils eine Woche vor und nach dem Kurztraining wurde in 30-45-minütigen Sitzungen die Professional Vision anhand von je zwei Videovignetten erfasst.

Das Kurztraining begann mit einem Einführungstext, der verschiedene Strategien einer Lehrkraft (z.B. einen kognitiven Konflikt hervorrufen) sowie deren Auswirkungen auf den Lernprozess der Schüler*innen vorstellte. Im Anschluss zeigte ein Beispielvideo eine Lehrkraft, die diese Strategien in einer Kleingruppensituation anwendete. Alle Studierenden sahen das Video zweimal und erhielten dabei entweder keine besondere Aufforderung für die beiden Durchgänge (Kontrollbedingung) oder sie wurden aufgefordert, im ersten Durchgang besonders auf das Verhalten der Lehrkraft zu achten und im zweiten Durchgang die Konsequenzen des Verhaltens für die Schüler*innen zu fokussieren (Emphasis-Shift-Training). Im Anschluss kommentierten die Studierenden das Video. Die Eingabemaske enthielt entweder offene Selbsterklärungsprompts oder fokussierte Prompts, die noch einmal an die Strategien aus dem Einführungstext erinnerten. Die Studierenden wurden jeweils randomisiert einer der beiden Bedingungen zugewiesen.

Die Videovignetten für den Prä- und Posttest basierten auf den gleichen Skripten und wurden mit zwei unterschiedlichen Gruppen von Schauspieler*innen gedreht. Bei der Bewertung der schriftlichen Kommentare der Studierenden wurden sowohl die Anzahl erkannter Ereignisse als auch für die Qualität der Beschreibung und Interpretation beachtet.

Ergebnisse und Diskussion

Im Posttest erkannten alle Probanden mehr unterschiedliche Ereignisse ($M = 3.20$, $SD = 1.24$) als im Prätest ($M = 2.40$, $SD = 1.27$), $F(1, 130) = 34.8$, $p < .001$, $\eta p2 = .21$. Außerdem waren ihre Beschreibungen und Interpretationen im Posttest von höherer Qualität, $F(1, 130) = 40.0$, $p < .001$, $\eta p2 = .24$. Entgegen der Erwartungen unterschieden sich die Bedingungen nicht signifikant im Ausmaß der Verbesserung ($F < 1$ für beide Interaktionen EST*Messzeitpunkt und Selbsterklärungsprompts*Messzeitpunkt). Eine Erklärung für diesen Befund könnte der große Effekt des Einführungstextes sein, der schon in früheren Studien mögliche kleinere Effekte anderer Maßnahmen während der Trainingsphase überlagert hat (Autoren, 2022). Außerdem ist die Dauer des Trainings im Vergleich zu anderen videobasierten Trainings zur Förderung der Professional Vision (z.B. Gold et al., 2021) eher kurz. Im Beitrag sollen Möglichkeiten und Grenzen eines Kurztrainings diskutiert werden.

Zentrale Befunde aus dem CLIPSS-Projekt mit Blick auf komplexitätsreduzierende Maßnahmen bei der Videoanalyse

Gerlinde Lenske[1], Julia Bönte[2], Rijana van Bebber[2], Detlev Leutner[2]
[1]Leuphana Universität Lüneburg, [2]Universität Duisburg-Essen

Videobasiertes Lernen ermöglicht eine Theorie-Praxis-Verzahnung bereits in der ersten Phase der Lehrkräftebildung, wobei das Medium Video als Fenster zur Praxis fungiert und die Praxis frei von Handlungsdruck erfahren und analysiert werden kann (Lenske, Bönte, van Bebber & Leutner, 2022). Evidenz für die Förderung professioneller Kompetenzen durch den Einsatz von Videos liegt vor (z.B. Gold, Förster & Holodynski, 2013; Gold, Pfirrmann & Holodynski, 2020; Sherin & van Es, 2009; Seidel & Thiel, 2017). Es zeigt sich jedoch auch, dass Videos für Noviz:innen sehr herausfordernd, teilweise sogar überfordernd sein können: Im Vergleich zu erfahrenen Lehrkräften fokussieren Noviz:innen häufiger unwesentliche Details (Santagata et al., 2007), registrieren seltener relevante Unterrichtsereignisse (Van den Bogert et al., 2014) und haben einen weniger flächendeckenden Blick auf das Unterrichtsgeschehen (Wolff, Jarodzka, van den Bogert & Boshuizen, 2016). Als Ursache wird die Komplexität der Unterrichtsvideos angenommen, mit zahlreichen simultanen Ereignissen auf verschiedenen Ebenen (Kramer et. al., 2017; Syring et al., 2015). Noviz:innen verfügen im Vergleich zu Expert:innen über ein geringeres Professionswissen und weniger ausgebaute kognitive Schemata (Van Gog, 2014). Dies reduziert die Kapazität, simultane Ereignisse kognitiv zu verarbeiten (vgl. Paas & Sweller, 2014). Nach der Cognitive Load Theory erschwert kognitive Überforderung (overload) die Informationsverarbeitung, was den Lernprozess beeinträchtigt. Bei der Analyse nonverbaler Indikatoren, welche beispielsweise im Bereich der Klassenführung von Bedeutung sind, können außerdem die Tonspur-Informationen die Verarbeitung nonverbaler Informationen erschweren: In Anlehnung an den unattended speech effect beeinträchtigt die simultan zum Bild ablaufende Tonspur die semantische Verarbeitung des Bildmaterials (Salamé & Baddeley, 1982).

Multimediale Präsentationstechniken wie Signaling (Hervorheben relevanter bzw. Ausblenden irrelevanter Ereignisse durch visuelle Signale zur Reduktion des extraneous cognitive load) und Segmenting (systematische Zerlegung von Lerninhalten in kleinere Einheiten zur Reduktion des intrinsic cognitive load) sind Möglichkeiten, die Videokomplexität zu reduzieren (Mayer, Moreno, Boire & Vagge, 1999; Mayer & Fiorella, 2014). Das Ausblenden der Tonspur kann zudem als Strategie betrachtet werden, um den unattended speech effect zu kompensieren (Reduktion des extraneous cognitive load). Dieser Beitrag prüft, inwiefern die genannten Strategien das Lernen anhand von klassenführungsbezogenen Unterrichtsvideos erleichtern.

In Studie I wurden anhand eines 2x2-experimentellen Designs (N=83 Lehramtsstudierende) die Faktoren Signaling (mit/ohne) und Segmenting (mit/ohne) variiert. Abhängige Variable war das anwendungsbezogene Klassenführungswissen (prä-post-Testung, gekürzte Testversion; Seethaler et al., 2021), Kontrollvariablen waren das situationale Interesse (FAM, sechs Items, neunstufige Likert-Skala, Rheinberg et al., 2001) und die kognitive Belastung (sechs Items, neunstufige Likert-Skala; Paas, 1992).

In Studie II wurde anhand eines 2-Gruppen-Designs (N=40 Lehramtsstudierende) das Ein- bzw. Ausblenden der Tonspur variiert, wobei in einer Bedingung das Video zweimal bei simultanem Abspielen von Ton- und Bildspur betrachtet wurde, während in der anderen Bedingung die Tonspur beim erstmaligen Betrachten ausgeblendet wurde. Abhängige Variable war das nonverbale Klassenführungswissen (prä-post-Testung, NCMS; Bönte, Lenske & Leutner, 2021).

In Studie I zeigte sich, dass das anwendungsbezogene Klassenführungswissen durch die Videointervention anstieg (d=0.56). Dabei erwies sich Signaling als lernförderlich (d=0.74), während Segmenting das situationale Interesse (d=0.82) und den Wissenserwerb reduzierte (d=0.61). Sowohl Signaling als auch das Segmenting verringerten die wahrgenommene Aufgabenschwierigkeit. In Studie II führte die Videointervention zu einem Anstieg des nonverbalen Klassenführungswissens (ohne Ton: d=0.68, mit Ton: d=0.53), wobei der Unterschied jedoch nicht signifikant war.

Die Ergebnisse zeigen eine das Arbeitsgedächtnis entlastende und lernförderliche Wirkung von Signaling, was beim Lernen mit Unterrichtsvideos insbesondere für Noviz:innen empfohlen werden kann. Segmenting hingegen führte zwar zu einer kognitiven Entlastung, jedoch ebenfalls zu reduziertem Interesse und reduziertem Lernzuwachs. Wir vermuten, dass die mögliche förderliche Wirkung durch zu starkes Segmenting verhindert wurde. In Bezug auf den unattended speech effect nehmen wir an, dass der Einsatz von Gruppendiskussionen den Effekt kompensiert haben könnte: Die Videos wurden nicht nur in Einzelarbeit, sondern ebenfalls in der Gruppe analysiert. Hier sind Folgestudien notwendig.

Kann eine ausgearbeitete Beispielanalyse die professionelle Wahrnehmung und das professionelle Wissen von Lehramtsstudierenden fördern?

Jasmin Lilian Bauersfeld, Bernadette Gold
Technische Universität Dortmund

Für die Umsetzung qualitätsvollen Unterrichts werden das professionelle Wissen (PW) und die professionelle Unterrichtswahrnehmung (PUW) als eine zentrale Voraussetzung gesehen (Blömeke et al., 2015). Ihre Förderung kann dazu beitragen, angehende Lehrkräfte möglichst früh auf

unterrichtliche Herausforderungen, wie z. B. der Klassenführung, vorzubereiten. Dafür hat sich die Analyse von Unterrichtsvideos als effektiv erwiesen (z.B. Stürmer et al., 2013). Allerdings haben Lehramtsstudierende aufgrund hoher kognitiven Belastung häufig Schwierigkeiten relevante Klassenführungsereignisse in komplexen Unterrichtssituationen überhaupt zu erkennen (Wolff et al., 2015) und benötigen demzufolge Anleitung, um von Unterrichtsvideoanalysen zu profitieren. Der Rückgriff auf empirisch gesicherte Befunde der Lehr-Lern-Forschung für die Gestaltung solcher angeleiteter Lernsettings trägt auch zu einer evidenzbasierten Lehrkräftebildung bei.

Demzufolge orientiert sich die aktuelle Studie an der Forschung zu ausgearbeiteten Lösungsbeispielen, die durch Selbsterklärungsprompts (und ggf. instruktionalen Erklärungen) zum Verstehen grundlegender Prinzipien beitragen und somit kognitive Ressourcen für die selbstständige Aufgabenbearbeitung bereitstellen (Renkl, 1997; van Gog et al., 2011). Ausgearbeitete Beispielanalysen (ABA) zu einem Unterrichtsvideo helfen Noviz*innen möglicherweise ebenso, Merkmale einer elaborierten Analyse (hier: Merkmale einer Expertenanalyse) zu verdeutlichen, bevor sie selbst eine Videoanalyse vornehmen. Solche Expertenmerkmale betreffen a) das Erkennen/Benennen von Tiefenmerkmalen der Unterrichtsqualität, b) die Fokussierung auf Kognitionen und Verhalten der Schüler*innen sowie c) Top-Down-Verarbeitungsprozesse (Berliner, 2001; Wolff et al., 2015).

Beispiele sind besonders lernförderlich, wenn sie Prompts zur Selbsterklärung (Chi et al., 1994) und - bei wenig Vorwissen instruktionale Erklärungen enthalten, wodurch Lehramtsstudierende unterstützt werden können, grundlegende Prinzipien für ihre eigene Analyse abzuleiten und mögliche fehlgeleitete Selbsterklärungen präventiv verhindert werden (Renkl, 2002). Häufig können Studierende diese grundlegenden Prinzipien einer ABA jedoch nicht unmittelbar auf ihre eigene Problembearbeitung übertragen (Renkl et al., 2004). Beim fading werden die instruktionalen Erklärungen und Prompts schrittweise ausgeblendet, was zu einer aktiven Auseinandersetzung mit der ABA anregen (Renkl et al., 2004) und die Übertragung der grundlegenden Prinzipien auf die eigene Analyse fördern soll.

Die vorliegende Studie verfolgte das Ziel, die Wirksamkeit von ausgearbeiteten Beispielanalysen (ABA) zu Unterrichtsvideoclips (mit und ohne fading) auf die PUW und das PW von Lehramtsstudierenden im Vergleich zu einer Kontrollgruppe und somit die Übertragbarkeit evidenzbasierter Lehr-Lern-Prinzipien auf die videobasierte Lehrkräftebildung zu untersuchen.

Methode

An der quasi-experimentellen Studie nahmen 406 Studierende an einem achtwöchigen videobasierten Seminar teil. Die Studierenden durchliefen vier Sitzungen zu Aspekten der Klassenführung. In jeder Sitzung lasen sie Texte zu verschiedenen Klassenführungsaspekten. Die Interventionsgruppen studierten zusätzlich verschiedene ABA. Die ABA-Gruppe (IG1; n=196) erhielt ABAs mit Selbsterklärungsprompts und instruktionalen Erklärungen zum Selbststudium. Die fading-Gruppe (IG2; n=148) ergänzte im Verlauf des Semesters immer mehr fehlende instruktionale Erklärungen und Prompts in den ABAs. Die Kontrollgruppe (n=62) durchlief das gleiche videobasierte Seminar, erhielt aber keine ABAs. Alle drei Gruppen analysierten daraufhin in jeder Sitzung

eigenständig zwei bis drei Videos mit Bezug zur Klassenführung. In einem Prä- und Post-Test wurde die PUW in einem Videotest (Cronbachs α=.86; Gold & Holodynski, 2017) und das PW mittels eines Wissenstest über Klassenführung (Cronbachs α=.41; Kurz & Lenske, i.V.) erhoben.

Ergebnisse und Diskussion

Eine ANOVA mit Messwiederholung zeigte, dass die Studierenden gruppenübergreifend ihre PUW, $F(1,404)=231.98$, $p<.001$, $\eta p^2=.365$, und ihr PW, $F(1,403)=23.12$, $p<.001$, $\eta p^2 = .054$, signifikant steigern konnten. In der PUW konnten Unterschiede in der Veränderung von Prä- zu Posttest zum Vorteil der Kontrollgruppe zu der IG1, $F(1,257)=5.80$, $p=.017$, $\eta p^2=.022$, aber keine Unterschiede zwischen der IG2 und der Kontrollgruppe, $F(1,208)=2.66$, $p=.105$, $\eta p^2=.013$, oder zwischen den Interventionsgruppen, $F(1,343)=0.63$, $p=.428$, $\eta p^2=.002$, gefunden werden. In Bezug auf das PW zeigten sich keine Unterschiede zwischen den drei Gruppen (KG vs. IG1: $F(1,256)=2.91$, $p=.089$, $\eta p^2=.011$; KG vs. IG2: $F(1,208)=2.43$, $p=.120$, $\eta p^2=.012$; IG1 vs. IG2: $F(1,342)=0.05$, $p=.819$, $\eta p^2<.001$).

Weitere Ergebnisse zu Prozessdaten, die Hinweise auf Lernprozesse der Lehramtsstudierenden geben, sowie Limitationen werden im Rahmen des Vortrags vorgestellt und diskutiert.

Minimalintervention zur Förderung der professionellen Wahrnehmung von Lehramtsstudierenden

Sylvia Gabel, Andreas Gegenfurtner
Universität Augsburg

Professionelle Wahrnehmung gilt als eine Schlüsselkompetenz im Lehrberuf (van Es & Sherin, 2010). Da die Interaktionen im Klassenzimmer komplex sind und oft simultan ablaufen, wird hohe Aufmerksamkeit und Professionalität von Lehrkräften eingefordert (Sherin & van Es, 2005). Vor allem in Bezug auf Klassenführung müssen Lehrer*innen problematische Situationen präzise und schnell identifizieren, um effektiv zu handeln (Emmer & Sabornie, 2015; Gold et al., 2021). Dabei zeigen Expert*innen-Noviz*innen-Unterschiede, dass erfahrene Lehrkräfte problematische Situationen schneller und akkurater wahrnehmen als wenig erfahrene Lehrkräfte (Grub et al., 2022; Wolff, 2016; Wolff et al., 2020).

Unterrichtsvideos können zur Ausbildung dieser Kompetenzen herangezogen werden (Gaudin & Chaliès, 2015; König et al., 2022; Sherin & van Es, 2005), wobei ein effektives Instruktionsformat wichtig ist: Durch die Aufgabenstellung sollten kognitive Schemata aktiviert werden und zu einer top-down Wahrnehmung führen (vgl. Gilboa & Marlatte, 2017; Pichert & Anderson, 1977; Pressley et al., 1992), wodurch die Aufmerksamkeit auf relevante Aspekte im Video gelenkt werden kann. Gleichzeitig ist die Befundlage dazu noch uneinheitlich. Beispielsweise konnten Grub et al. (2022) keinen signifikanten Unterschied bezüglich einer spezifischen und allgemeinen Aufgabenstellung feststellen, wobei als Limitation die unzureichende Spezifität der Aufgabenstellung genannt wurde. Martin et al. (2022) erwarteten, dass Situationen akkurater und qualitativ hochwertiger erklärt werden, wenn fokussierte Prompts anstelle von offenen Prompts gezeigt werden, was jedoch nicht bestätigt werden konnte. Die Autor*innen konkludieren, dass eine spezifische Aufgabenstellung und die Darbietung von Prompts einen ähnlichen kognitiven Effekt auf die Aktivierung von relevanten

Schematat haben und dadurch einen aufmerksamkeitslenkenden Effekt durch den top-down Prozess hervorrufen können.

Fragestellung

Basierend auf der zugrundeliegenden Forschung zur Förderung von professioneller Wahrnehmung bezieht sich diese Studie auf folgende Fragen: Gibt es einen Unterschied zwischen (a) der Bereitstellung einer spezifischen Aufgabenstellung vor Videobetrachtung oder (b) der Darbietung von Prompts während der Videobetrachtung bezüglich der Anzahl der relevanten Aspekte (H1), der Relevanz der identifizierten Situationen (H2) und der Wahrnehmungsgeschwindigkeit (H3)?

Methode

Insgesamt wurden N = 85 Lehramtsstudierende (66 weiblich) der Universität Augsburg (MAlter = 21.29; SDAlter = 2.87) randomisiert in zwei Gruppen aufgeteilt (Spezifische Aufgabenstellung N = 42; Prompts N = 43). Das Mixed-Method-Design trianguliert die Klick-Aufzeichnungen der Studierenden während der Videobetrachtung und der Erklärungen im anschließenden Interview, weshalb sie die Situationen als relevant für die Klassenführung empfunden haben. Abschließend wurden in einem Fragebogen die mentale Anstrengung (Paas, 1992), das pädagogisch-psychologische Wissen (Lenske et al., 2015) sowie demographische Daten abgefragt.

Ergebnisse

Da die Gruppen keine signifikanten Unterschiede bezüglich der mentalen Anstrengung und des deklarativen pädagogisch-psychologischen Wissens aufweisen, können diesbezüglich unerwünschte Einflusseffekte ausgeschlossen werden.

Die Analysen zeigten keinen statistisch signifikanten Unterschied in der Anzahl der relevanten Situationen (H1) ($t(83) = -1.10$, $p = .28$), der Relevanz der Wahrnehmung (H2) ($t(82) = = -0.56$, $p = .57$) und der Wahrnehmungsgeschwindigkeit (H3) ($t(82) = -0.56$, $p = .57$) zwischen den beiden Gruppen (vgl. Tab. 1).

Diskussion

Für die Förderung der professionellen Wahrnehmung sind effektive Lernumgebungen in der Lehrkraftausbildung essenziell (König et al., 2022). Die Ergebnisse dieser Studie deuten darauf hin, dass die spezifische Aufgabenstellung einen vergleichbaren Effekt auf die Geschwindigkeit, die Relevanz und die Anzahl an Wahrnehmungen hat wie die Darbietung von Prompts. Auch die mentale Anstrengung wurde in beiden Gruppen als ähnlich hoch wahrgenommen. Eine praktische Implikation ergibt sich für den universitären Kontext: Die Darbietung von spezifischen Aufgabenstellungen ist eine ökonomischere Instruktionsform für den Einsatz von Unterrichtsvideos in Seminaren. Zukünftig kann auch Eye Tracking bei der Untersuchung von Minimalinterventionen eingesetzt werden (vgl. Cortina et al., 2015). Interessant wäre insbesondere eine Analyse der Fixationsmuster in Abhängigkeit zum Instruktionsformat.

C-14: Psychische Gesundheit von Schüler:innen als Dimension inklusiver Schule: Beiträge aus der Lehrkräfteprofessionalisierungsforschung

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* T03 R02 D81

Psychische Gesundheit von Schüler:innen als Dimension inklusiver Schule: Beiträge aus der Lehrkräfteprofessionalisierungsforschung

Chair(s): **Franziska Greiner** (Leuphana Universität Lüneburg, Deutschland), **Poldi Kuhl** (Leuphana Universität Lüneburg)
*Diskutant*in(nen):* **Orkan Okan** (Technische Universität München)

Schulische Inklusion im weiten Verständnis impliziert, dass alle Kinder und Jugendlichen die Möglichkeit erhalten, sich bestmöglich zu entwickeln. Obwohl jede:r vierte Schüler:in unter einer psychischen Störung leidet (Greiner, Batram & Witte 2019), wurde diese Zielgruppe in der empirischen Bildungsforschung bislang nicht explizit fokussiert. Daher liegen im deutschsprachigen Raum sowohl zur schulischen Situation von Kindern und Jugendlichen mit psychischen Störungen als auch zur Professionalisierung von Lehrkräften für diese Thematik kaum Befunde vor.

An dieses Desiderat anknüpfend, bündelt das Symposium Beiträge, die sich der empirischen Forschung zum professionellen Umgang von Lehrkräften mit psychischen Störungen von Schüler:innen widmen. Dabei werden verschiedene Ebenen betrachtet: Lehrkräfte, Schulleitungen und die Schule als Organisation.

Beitrag I widmet sich einer Fortbildung für Lehrkräfte, die auf die Förderung der Kompetenzen im Umgang mit Depressionen bei Schüler:innen abzielt. Die Fortbildung wurde mit einem Prä-Post-Follow-up-Design evaluiert. Dabei zeigten sich im Post- und Follow-up-Test sowohl signifikante Anstiege der Depressionskompetenz, des Handlungswissens als auch der subjektiven Handlungskompetenz.

In Beitrag II wird der professionelle Umgang von Lehrkräften mit nicht-suizidalen selbstverletzenden Verhaltensweisen (NSSV) von Schüler:innen beleuchtet. Ausgehend von einer textbasierten Fallvignette benannten die teilnehmenden Lehrkräfte Warnhinweise für NSSV, formulierten Handlungsschritte und gaben notwendige Unterstützungssysteme an. Nur zwei Lehrkräfte konnten alle in der Fallvignette zu findenden Hinweise auf NSSV identifizieren. Der Großteil würde zunächst ein persönliches Gespräch führen und Kontakt zur Schulsozialarbeit aufnehmen. Die Ergebnisse weisen darauf hin, dass die Lehrkräfte nur über oberflächliches Wissen in Bezug auf NSSV verfügen, welches sie nicht zu einem professionellen Umgang mit Schüler:innen befähigt, die sich selbst verletzen.

In **Beitrag III** wird die Perspektive von Schulleitungen hinsichtlich der Unterstützungsbedarfe von Schüler:innen im Bereich der psychischen Gesundheit fokussiert. Dafür werden Daten aus der COVID-HL-Onlinestudie analysiert. Angenommen wird, dass weibliche Schulleitungen, Grundschulleitungen und Schulleitungen aus Schulen mit höherem Anteil von Schüler:innen aus Familien mit geringem Sozialstatus einen höheren gesundheitlichen Unterstützungsbedarf der Schüler:innen wahrnehmen.

In **Beitrag IV** werden schließlich Zusammenhänge zwischen dem Maß an organisationaler Gesundheitskompetenz, der Gesundheitskompetenz der Schulleitungen und dem Umsetzungsstand von Maßnahmen schulischer Gesundheitskompetenz- und Gesundheitsförderung an deutschen Schulen beleuchtet. Die Daten werden bis November 2022 erhoben. Zu erwarten ist, dass eine

höhere organisationale Gesundheitskompetenz der Schule positiv auf die psychische Gesundheitskompetenz der Schüler:innen und die Gesundheitskompetenz aller Akteur:innen wirkt.

Da der professionelle Umgang mit psychischen Störungen als eine Facette inklusiver Lehrkräftekompetenzen verstanden wird, werden die Beiträge des Symposiums vor dem Hintergrund der Anforderungen an ein inklusives Schulsystem diskutiert.

Beiträge des Symposiums

Depressionen im Kindes- und Jugendalter - Entwicklung und Evaluation einer psychoedukativen Fortbildung für Lehrkräfte

Friederike-Carlotta Grabowski
Europa-Universität Flensburg

Depressive Störungen sind eine der häufigsten psychischen Erkrankungen im Kindes- und Jugendalter weltweit (Shorey et al., 2022) mit einer Prävalenz von 5-10 % in Deutschland (Groen & Petermann, 2013; Schulte-Körne, 2018; Wartberg et al., 2018). Die Coronavirus Pandemie 2019 (COVID-19) hat die psychische Belastung in dieser Altersgruppe erhöht und das Problem psychischer Störungen verschärft (Marques de Miranda et al., 2020; Ravens-Sieberer et al., 2021). Früh einsetzende depressive Erkrankungen führen zu den langwierigsten Beeinträchtigungen im Leben, einschließlich eines erhöhten Suizidrisikos (Petermann & Vries, 2016) sowie sozialen, gesellschaftlichen und wirtschaftlichen Nachteilen (Essau, 2007; Gander & Buchheim, 2013). Obwohl Lehrkräften als Bezugspersonen eine zentrale Rolle in der Früherkennung und Hilfevermittlung depressiver Schülerinnen und Schüler zukommt, fehlt es an geeigneten psychoedukativen Fortbildungen zur Steigerung der Depressionskompetenz und des Handlungswissens. Gegenwärtig bleiben depressive Symptome von Schülerinnen und Schülern daher oft unerkannt durch Lehrkräfte (Cunningham & Suldo, 2014; Gander & Buchheim, 2013; Splett et al., 2019). Dies trägt unter anderem dazu bei, dass Betroffene häufig nicht oder erst spät behandelt werden (Klasen et al., 2017; Ravens-Sieberer et al., 2007; Wittchen et al., 2010).

Um diese Situation zu verbessern, wurde ein zweistündiges psychoedukatives Trainingsprogramm für Lehrkräfte namens glücklich entwickelt und pilotiert. Glücklich vermittelt Wissen über Depressionen im Kindes- und Jugendalter mit einem Fokus auf den schulischen Bereich in einem zeitökonomischen Rahmen. Neben theoretischem Wissen wird Handlungswissen anhand von Fallbeispielen nähergebracht. In ersten Studien wurde untersucht, inwieweit sich die Depressionskompetenz, das Handlungswissen und die subjektive Handlungskompetenz von Lehrkräften durch das Training verändern.

In einer formativen und zwei summativen Studien wurde glücklich online mit Lehramtsstudierenden (n = 84) und Lehrkräften (n = 7) evaluiert (Grabowski, 2021). In beiden summativen Studien wurde das Programm in einem Prä-Post-Follow-up-Design untersucht. Da kein geeignetes standardisiertes Messinstrument bezüglich Depressionskompetenz, Handlungswissen und subjektiver Handlungskompetenz existiert (Yamaguchi et al., 2020), wurde ein eigener Fragebogen entwickelt. Teilnehmende beantworteten 24 Fragen bezüglich der Depressionskompetenz (z. B. „Depressionen sind eine behandelbare medizinische Krankheit“) auf einer zweistufigen Skala („stimme zu“/ „stimme nicht zu“). Zur subjektiven Handlungskompetenz (z. B. „Ich fühle mich kompetent im Erkennen

depressiver Symptome bei Schüler:innen“) konnten vier Fragen auf einer vierstufigen Skala („stimme überhaupt nicht zu“ bis „stimme voll und ganz zu“) beantwortet werden. Mit 11 bzw. 12 Fragen in Bezug auf zwei Fallvignetten (z. B. „Den schulpsychologischen Dienst zu Rate ziehen“) wurde das Handlungswissen auf einer dreistufigen Skala („hilfreich“/ „nicht hilfreich“/ „kann ich nicht beurteilen“) erfasst.

In Studie 1 (n1 = 8) beantworteten Lehramtsstudierende nach der Fortbildung einen formativen Fragebogen. Dieser ergab, dass einige Verbesserungen erforderlich waren. In Studien 2 und 3 nahmen n = 91 (n2 = 78; n3 = 13) Lehramtsstudierende und Lehrkräfte an der Fortbildung und Befragung teil (Ausgangsstichprobe: n2 = 113; n3 = 32). Ein signifikanter ($p \leq .05$) Anstieg der Depressionskompetenz (Studie 2: $\eta 2 = 0.512$; Studie 3: $\eta 2 = 0.664$) und des Handlungswissens (Studie 2: $\eta 2 = 0.229$ / $\eta 2 = 0.353$; Studie 3: $\eta 2 = 0.597$ / $\eta 2 = 0.609$) wurde im Posttest beobachtet und im Follow-up nach vier Wochen beibehalten. Auch hinsichtlich ihrer subjektiven Handlungskompetenz (Studie 2: $\eta 2 = 0.44–0.66$; Studie 3: $\eta 2 = 0.50–0.80$) zeigten die Teilnehmenden signifikante Verbesserungen im Posttest, welche im Follow-up Bestand hatten (Grabowski, 2021). Die Pilotierung von glücklich zeigt vielversprechende vorläufige Ergebnisse für die Verbesserung der Früherkennung von depressiven Schülerinnen und Schülern und der Fähigkeit von Lehrerkräften, diesen angemessen zu helfen. Es konnte insbesondere gezeigt werden, dass auch mit einem kurzen Programm wie glücklich signifikante und zeitstabile Wissenszunahmen von Lehrkräften erzielt werden können, wovon viele depressive Schülerinnen und Schüler profitieren würden.

Handlungsbezogenes Wissen von Lehrkräften über nicht-suizidales selbstverletzendes Verhalten von Schüler:innen

Franziska Greiner[1], **Bärbel Kracke**[2]
[1]Leuphana Universität Lüneburg, [2]Friedrich-Schiller-Universität Jena

Hintergrund. Nicht-suizidales selbstverletzendes Verhalten, welches als direkte und vorsätzliche Körperverletzung ohne tödliche Absicht verstanden werden kann, stellt ein ernstes und wachsendes psychisches Gesundheitsproblem bei Jugendlichen im Schulalter dar (Hamza & Heath, 2018). Innerhalb Europas gehört Deutschland zu den Ländern mit den höchsten Prävalenzraten (Brunner et al., 2014; Plener et al., 2009).

Dass Lehrkräfte an allen Schulformen mit nicht-suizidalem selbstverletzendem Verhalten (NSSV) von Schüler:innen konfrontiert werden, ist sehr wahrscheinlich. Trotz der hohen Prävalenz, der zusätzlichen psycho-emotionalen Belastung der Schüler:innen durch die COVID-19-Pandemie (Ravens-Sieberer et al., 2020) und der zentralen Rolle von Lehrkräften in Bezug auf die frühzeitige Identifizierung von psychischen Auffälligkeiten und der Vermittlung von professioneller Unterstützung (Greiner & Kracke, 2022) konnte im deutschsprachigen Raum keine Studie zum Umgang von Lehrkräften mit NSSV von Schüler:innen gefunden werden.

Dieses Desiderat greift der vorliegende empirische Beitrag auf und untersucht in einer explorativen Studie, über welches handlungsbezogene Wissen Lehrkräfte weiterführender Schulen im Umgang mit NSSV bei Schüler:innen verfügen.

Methode und Stichprobe. Eingesetzt wurde ein online-Fragebogen mit überwiegend offenen Items, die auf eine textbasierte Fallvignette rekurrieren. Die Fallvignette umfasst acht Hinweise auf NSSV

(z.B. sozialer Rückzug, Sportatteste, langärmlige Kleidung). Die Lehrkräfte wurden nach Rezeption der Fallvignette darum gebeten, die Warnhinweise für NSSV zu benennen, anschließend Handlungsschritte zu formulieren sowie notwendige Unterstützungssysteme anzugeben.

Insgesamt wurden die Daten von 121 Lehrkräften (♀: 102, ♂: 17, k.A.: 2) weiterführender Schulen durch zwei Forscherinnen unabhängig voneinander codiert und quantifizierend ausgewertet. Das mittlere Alter der befragten Lehrkräfte betrug 46.1 Jahre (SD = 11.74). Im Durchschnitt hatten die Lehrkräfte 20 Jahre Berufserfahrung (min = 0.25; max = 45).

Ergebnisse und ihre Bedeutung. 79 % der befragten Lehrkräfte berichteten von Erfahrungen mit NSSV von Schüler:innen, von denen knapp 80 % in ihrem beruflichen Alltag bereits mit mehr als einer Form von NSSV konfrontiert wurden. Nur zwei der 121 Lehrkräfte waren in der Lage, alle acht Hinweise auf NSSV zu nennen. Der Großteil der Befragten konnte zwei oder drei Hinweise für NSSV identifizieren. Am häufigsten wurden die Schnitte an den Unterarmen erkannt (71 %), gefolgt von sozialem Rückzug (63 %) und langärmliger Kleidung (55 %). Seltener wurden die Sportatteste (27 %) und die verfassten, traurigen Texte (25 %) als mögliche Hinweise auf NSSV identifiziert. Kaum erkannt wurden der Leistungsabfall sowie Annas Angespanntheit und mentale Abwesenheit im Unterricht.

Bei der Bearbeitung der Fallvignette zeigte sich, dass 95 % der Lehrkräfte zunächst ein persönliches Gespräch mit der betroffenen Schülerin suchen würden. Etwas weniger als die Hälfte würde sofort oder im Anschluss an das Gespräch die Sorgeberechtigten informieren. 12,5 % würden Mitschüler:innen einbeziehen, 38 % würden sich an Kolleg:innen bzw. Schulsozialarbeiter:innen und Schulpsycholog:innen wenden und 9 % würden sich externe Beratung einholen, bevor weitere Handlungsschritte erfolgen.

Insgesamt wird deutlich, dass die meisten Lehrkräfte einen ersten Kontakt mit der betroffenen Schülerin aus der Fallvignette aufnehmen würden. Jedoch zeigen sie große Unsicherheiten im weiteren Umgang mit der Schülerin und den im Gespräch gewonnenen Informationen.

Der vorliegende Forschungsbeitrag verdeutlicht die Notwendigkeit einer Sensibilisierung von Lehrkräften für NSSV von Schüler:innen und den damit einhergehenden Bedarf einer Professionalisierung, die Lehrkräfte dazu befähigt, konkrete Handlungsschritte im Umgang mit betroffenen Schüler:innen professionell zu planen und umzusetzen. Vor dem Hintergrund der gewonnenen Erkenntnisse werden Möglichkeiten eines Professionalisierungsangebots in der Lehrkräfteaus-, Fort- und -weiterbildung diskutiert.

Gesundheitliche Unterstützungsbedarfe von Schüler*innen aus Sicht von Schulleitungen. Ergebnisse der COVID-HL Schulleitungsstudie

Kevin Dadaczynski[1], Melanie Messer[2], Marlene Meyer[3], Orkan Okan[3]
[1]Hochschule Fulda, [2]Universität Trier, [3]Technische Universität München

Hintergrund: Repräsentative Surveys weisen darauf hin, dass etwa 17% aller Kinder und Jugendlichen bis 17 Jahren psychische Auffälligkeiten aufweisen, wobei Mädchen und Jungen aus Familien mit niedrigem sozioökonomischen Status deutlich häufiger betroffen sind (Klipker et al., 2018). Die gesundheitliche Situation dürfte sich infolge der seit Frühjahr 2020 bestehenden Corona-Pandemie noch einmal deutlich verschärft haben. So geben 71% der im Rahmen der COPSY Studie

befragten 7- bis 17-Jährigen Kinder und Jugendlichen an, sich durch die erste Welle der Pandemie belastet zu fühlen (Ravens-Sieberer et al., 2021). Dabei hat sich der Anteil an psychischen Auffälligkeiten im Vergleich zu der Zeit vor der Pandemie verdoppelt (ebd.). Angesichts gestiegener psychischer Belastungen und Beschwerden wächst die Bedeutung der Schule als ein zentrales und historisch gewachsenes Setting der Gesundheitsförderung und Prävention. Ob und in welchem Ausmaß Maßnahmen der schulischen Gesundheitsförderung umgesetzt werden, hängt dabei maßgeblich von Wahrnehmungen, Einstellungen und Gesundheitskompetenzen der Schulleitungen ab (Dadaczynski et al., 2020). Dabei messen Schulleitungen externalisierenden gegenüber internalisierenden Problemen von Schüler*innen eine deutlich höhere Bedeutung bei (Dadaczynski & Paulus, 2011).

Fragestellung: Infolge der sehr begrenzten Studienlage zielt die vorliegende Untersuchung auf die Wahrnehmung gesundheitlichen Unterstützungsbedarfe von Schüler*innen aus Sicht von Schulleitungen im Kontext der Corona-Pandemie. Dabei sollen Unterschiede der Wahrnehmung in Abhängigkeit individueller Faktoren, der Schulform und des Anteils von Schüler*innen aus geringem sozioökonomischen Status ermittelt werden. Zudem wird der Frage nachgegangen, ob ein höher wahrgenommener gesundheitlicher Bedarf mit einem höheren Umsetzungsstand von Maßnahmen der schulischen Gesundheitsförderung assoziiert ist.

Methoden: Es werden Daten der COVID-HL Onlinestudie herangezogen, die von März bis April 2021 mit 2186 Schulleitungen aus Baden-Württemberg, Hessen, Niedersachsen und Nordrhein-Westfalen durchgeführt wurde. Die Erfassung gesundheitlicher Bedarfe von Schüler*innen erfolgte über acht Items, die von Schulleitungen auf einer vierstufigen Skala (sehr geringe bis sehr hohe Bedeutung) beantwortet werden konnten. Als individuelle Faktoren wird neben dem Alter (offen erfasst) und dem Geschlecht (männlich, weiblich, divers) die coronaspezifische Gesundheitskompetenz (Okan et al., 2022) herangezogen. Die Erfassung des sozioökonomischen Status erfolgte über eine Einschätzung des prozentualen Anteils der Schüler*innen aus Familien der höheren, mittleren oder niedrigen sozialen Schicht. Der Umsetzungsstand der schulischen Gesundheitsförderung wurde auf Basis von Vorarbeiten mittels 15 Items operationalisiert (Dadaczynski et al., 2020). Die Analyse der Daten erfolgt mittels Unterschiedstests sowie multiplen linearen Regressionsanalysen.

Ergebnisse und ihre Bedeutung: Die Daten befinden sich derzeit in der Auswertung, wobei davon ausgegangen wird, dass weibliche Schulleitungen und Grundschulleitungen eine höhere Sensibilität aufweisen und entsprechend einen höheren gesundheitlichen Unterstützungsbedarf für Schüler*innen angeben. Ebenso wird davon ausgegangen, dass Schulleitungen aus Schulen mit höherem Anteil von Schüler*innen aus Familien mit geringem Sozialstatus und jene mit einer höheren Gesundheitskompetenz einen höheren gesundheitlichen Unterstützungsbedarf benennen. Zum Zusammenhang zwischen gesundheitlichem Unterstützungsbedarf für Schüler*innen und dem Umsetzungsstand der schulischen Gesundheitsförderung werden in Ermangelung an Vorbefunden keine Hypothesen formuliert. Die Ergebnisse der Untersuchungen versprechen wichtige Impulse für die schulische Gesundheitsförderung, so u. a. für die Entwicklung von Maßnahmen zur Sensibilitätssteigerung und zur Förderung der Gesundheitskompetenz bei Schulleitungen.

Organisationsentwicklung der Schule und Verhältnisprävention: Erstmalige Messung der organisationalen Gesundheitskompetenz von Schulen

Cara Krudewig, Sandra Kirchhoff, Orkan Okan
Technische Universität München

Hintergrund: Ein gesundes Aufwachsen von Kindern und Jugendlichen erfordert insbesondere die Schaffung von Rahmenbedingungen, die systematisch auf soziale, kulturelle und ökonomische Herkunftsunterschiede eingehen (Bauer & Bittlingmayer, 2015). Die Daten des KiGGs Surveys des Robert Koch-Instituts belegen seit mehr als einer Dekade, dass die Gesundheit im Lebensverlauf und das Gesundheitsverhalten, die Einstellungen und Kompetenzen von Kindern und Jugendlichen einem sozialen Gradienten folgen (Kuntz et al., 2018). Häufig verhindert das Präventionsdilemma allerdings, dass Maßnahmen benachteiligte Kinder und Jugendliche erreichen (Bauer, 2005). Schule als Bildungsort wird in diesem Zusammenhang häufig als ideales, nachhaltiges Interventionssetting benannt (Paulus, 2022; Dadaczynski et al., 2022). Zum einen weil dort alle schulpflichtigen Kinder über einen langen Zeitraum niedrigschwellig erreicht werden können und zum anderen weil Schule eine Prävention erlaubt, die gesundheitsbezogenen Dispositionen früh adressiert(Okan et al 2020). Mit Gesundheitskompetenz (eng. Health Literacy) existiert seit einigen Jahren ein Konzept, dass die schulische Gesundheitsförderung und -bildung um die Komponente gesundheitsbezogener Informationsverarbeitung in digitalen und analogen Kontexten erweitert (Schulenkorf et al., 2022). Mit dem Ansatz der organisationalen Gesundheitskompetenz von Schulen wird im GeKoOrg-Schule Projekt ein ganzheitlicher, inklusiver und verhältnispräventiver Ansatz entwickelt und untersucht, um über Strukturen und Akteur:innen im Bildungskontext die Gesundheit von Kindern und Jugendlichen zu verbessern (Kirchhoff & Okan, 2022; Kirchhoff et al., 2022a/2022b).

Fragestellung: Im Anschluss an eine partizipative Konzept- und Fragebogenentwicklung mit Lehrkräften, Schulleitungen und Forschenden wird gegenwärtig ein Online-Survey zur organisationalen Gesundheitskompetenz (oGK) von Schulen in NRW durchgeführt. Im Rahmen der Studie werden Schulleitungen als zentrale Akteur:innen schulischer Gesundheitsförderung befragt mit dem Ziel, erstmalig die oGK von Schulen zu erfassen und Zusammenhänge zwischen dem Maß an oGK, der Schulleiter-bezogenen Gesundheitskompetenz und dem Umsetzungsstand von Maßnahmen schulischer Gesundheitskompetenz- und Gesundheitsförderung an deutschen Schulen zu untersuchen.

Methoden: Die oGK der Schule wird über eine neuentwickelte Skala (Okan & Kirchhoff, 2022) erfasst, in der acht Standards einer “gesundheitskompetenten“ Schule mit jeweils sechs Indikatoren (48 Items) enthalten sind. Über die Kurzform des HLS19 (HLS-EU-Q12, Sørensen et al., 2013) wird die Gesundheitskompetenz der Schulleitung erfasst. Zusätzlich werden soziodemografische Merkmale der Schulleitungen und der Schule (z.B. Anzahl der Schüler:innen, Schulform) sowie weitere validierte Skalen zur Gesundheit an der Schule und der Schulleitungen erhoben. Im Anschluss an die onlinebasierte Querschnitterhebung im Herbst 2022 werden die Daten zunächst deskriptiv berichtet und mittels weiterer inferenzstatistischer Methoden ausgewertet.

Ergebnisse und ihre Bedeutung: Die Daten werden derzeit erhoben und ab Dezember 2022/Januar 2023 ausgewertet. Derzeit liegt keine empirische Evidenz zur oGK der Schule vor. Studien in anderen Settings (Versorgung, Jugendarbeit) sowie zur Gesundheitskompetenz von Schulleitungen belegen

allerdings, dass der Ansatz über die Schulstrukturen und Rahmenbedingungen positive Effekte auf die Umsetzung von gesundheitlichen Maßnahmen in den jeweiligen Settings hat (Dadaczynski et al., 2020; Okan et al., 2021). Es ist zu erwarten, dass Schulen das derzeit noch relativ neue Konzept der Gesundheitskompetenz noch nicht umsetzen, was auch daran liegt, dass Schulen kein Mandat für die Umsetzung von Gesundheit haben und schulische Gesundheitsförderung eher freiwillig durch Lehrkräfte oder Schulleitungen in das Setting gebracht werden. Aus verhältnispräventiver Sicht ist zu erwarten, dass eine höhere oGK der Schule positiv auf die Gesundheitskompetenz der Schüler:innen und die Gesundheitskompetenz aller im Setting beteiligten Personen wirkt. Zudem sollte die Gesundheitskompetenz der Schulleitung positiv auf die oGK der Schule wirken.

C-15: Trainingsprogramme zur sprachlichen Förderung im Unterricht: Möglichkeiten und Grenzen

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* T03 R03 D75

Trainingsprogramme zur sprachlichen Förderung im Unterricht: Möglichkeiten und Grenzen

Chair(s): **Simone Jambor-Fahlen** (Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache, Universität zu Köln)
*Diskutant*in(nen):* **Elmar Souvignier** (Institut für Psychologie in Bildung und Erziehung, Westfälische Wilhelms-Universität Münster)

Zur wirksamen Förderung schriftsprachlicher Fähigkeiten liegen zahlreiche Erkenntnisse aus empirischen Studien vor (Graham & Harris, 2017; Graham et al. 2017; Philipp, 2013). Darauf aufbauend setzen einzelne Bundesländer Programme auf, um das Lesen und Schreiben flächendeckend an ihren Schulen zu fördern: so zum Beispiel *Die Textprofis* in Baden-Württemberg, *FILBY* in Bayern oder *Die Schreibstarken* in Hamburg und Berlin. Gemeinsam ist diesen Programmen, dass sie trainingsförmige Konzepte beinhalten. Trainingsprogramme sind dadurch gekennzeichnet, dass strukturierte, aufeinander bezogene Aufgaben planmäßig und nicht beiläufig durchgeführt werden, um klar definierte Kompetenzen systematisch zu fördern. Sie eignen sich somit gut für das regelmäßige und wiederholende Üben. Von Vorteil für Schüler:innen ist zudem der transparente Zusammenhang von Training und Lernfortschritt. Bestehend aus Trainingsmanual und Aufgabenpaketen zeichnen sich Trainingsprogramme dadurch aus, dass sie „zugleich Träger theoretischer Prinzipien sind und praktische Umsetzungsvorschläge enthalten“ (Souvignier & Ehlert, 2022). Die Wirksamkeit solcher Trainingsprogramme hängt stark von der konzepttreuen Umsetzung ab. In der schulischen Wirklichkeit – außerhalb von kontrollierten Interventionen – werden Trainingsprogramme von Lehrkräften durchgeführt, die wiederum von Multiplikator:innen oder Wissenschaftler:innen qualifiziert wurden. Hier entsteht eine „Filterkette“ (Roth & Skintey, 2022, S. 79), die sich auf die konzeptgemäße Durchführung auswirken kann. So zeigen Studien, dass der Einsatz von Förderprogrammen in der Praxis oftmals geringere Effektstärken hinsichtlich der Wirksamkeit aufweist als in kontrollierten Evaluationsstudien (Schneider, 2019). Der Begriff *research-to-practice-gap* (Souvignier & Philipp, 2016) verweist auf diese Diskrepanz.

Das Symposium greift diesen Umstand auf und geht der Frage nach, wie sich der Zusammenhang zwischen der Umsetzung systematischer Trainingsprogramme in der Schulpraxis und deren Wirksamkeit gestaltet. Eng verbunden hiermit sind Fragen nach Durchführungsbedingungen, Durchführungstreue, Differenzierung und Akzeptanz. Ziel des Symposions ist es, Gelingensbedingungen und Hindernisse bei der Implementation von Trainingsprogrammen zu identifizieren, die möglichst früh bei der Entwicklung von Förderkonzepten berücksichtigt werden müssen, um so zu einer Verringerung der *research-to-practice-gap* beizutragen und eine langfristige Implementation in die Fläche zu ermöglichen. Die Beiträge diskutieren die Fragestellung unter verschiedenen Perspektiven:

Der erste Beitrag diskutiert die differenzielle Wirksamkeit von Schreibflüssigkeitstrainings vor dem Hintergrund heterogener Schulklassen und innerer Differenzierung.

Der zweite Beitrag eruiert, wie das Lesestrategietraining FiLBY-3 im Präsenzunterricht und im pandemiebedingten Distanzunterricht von Lehrkräften durchgeführt wurde und leitet daraus Bedingungen für einen erfolgreichen Transfer ab.

Der dritte Beitrag identifiziert Erfolgsfaktoren und Herausforderungen bei der Umsetzung eines Schreibstrategietrainings abgeleitet aus der Perspektive von Lehrkräften, Schüler:innen und Fachberater:innen.

Der Diskussionsbeitrag stellt die aus Sicht der Sprachdidaktik eruierten Implementationsfaktoren für Trainingsprogramme zur sprachlichen Förderung in den Kontext der psychologischen Schulevaluationsforschung.

Beiträge des Symposiums

Differenzielle Effekte von Schreibflüssigkeitstrainings – heterogene Lerngruppen und schulische Rahmenbedingungen

Sabine Stephany[1], Anne Griepentrog[2], Alina Koch[2], Valerie Lemke[3], Michael Becker-Mrotzek[2]
[1]Institut für deutsche Sprache und Literatur II, Universität zu Köln, [2]Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache, Universität zu Köln, [3]Institut für Germanistik, Universität Duisburg-Essen

Texte schreiben zu können gehört zu den Schlüsselkompetenzen in unserer Gesellschaft. Eine wesentliche Voraussetzung für eine entwickelte Schreibfähigkeit ist das mühelose Beherrschen der basalen Schreibfertigkeiten, insbesondere der Schreibflüssigkeit. Flüssiges Schreiben wird jedoch nicht von allen Schüler:innen in Primar- und Sekundarstufe in ausreichendem Maße beherrscht, so dass sie viele Aufgaben im Unterricht aller Fächer nicht bewältigen können. Schreibflüssigkeit setzt sich aus dem schnellen Formulieren kurzer Wortverbindungen, also dem mentalen „Versprachlichen“ von textrelevanten Ideen (Formulierungsflüssigkeit), und dem schnellen, automatisierten und korrekten Aufschreiben dieser Wörter (Transkriptionsflüssigkeit) zusammen (Linnemann et al., in Dr.; McCutchen, 1996). Sind diese Fertigkeiten noch nicht automatisiert, stehen nicht genügend kognitive Ressourcen für hierarchiehohe Prozesse zur Verfügung wie das Planen oder die Adressatenantizipation. Unflüssiges Schreiben wirkt sich daher negativ auf die Textqualität aus (Alves & Limpo, 2015; Chenoweth & Hayes, 2001).

Trotz ihrer Relevanz steht die Schreibflüssigkeit bisher kaum im Fokus schreibdidaktischer Forschung und Unterrichtspraxis. Es gibt kaum empirisch erprobte Förderkonzepte. Daher stellt sich die Frage, wie Schreibflüssigkeitstrainings konzipiert und in Abhängigkeit von schulischen Rahmenbedingungen wirksam in den Unterricht implementiert werden können. Zur Beantwortung der Fragestellung wurde in einer ersten Interventionsstudie ein Schreibflüssigkeitstraining entwickelt. Die Materialen wurden in Anlehnung an ein Sporttraining konzipiert: Sie sollen motivieren und den Schüler:innen verdeutlichen, dass flüssiges Schreiben trainier- und erlernbar ist. Das Trainingsmaterial umfasst Aufgaben zum automatisierten Abruf von Wörtern und Wortverbindungen, wie Abschreibaufgaben und das Ergänzen von typischen Wortkombinationen, sowie Aufgaben zum flüssigen Formulieren, wie das Erweitern von Phrasen, das Verbinden von Sätzen mit Konnektoren oder das Assoziieren zu Begriffen. Das Training besteht aus 18 Einheiten à 15 Minuten, die in einem Prä-Post-Kontrollgruppendesign zweimal pro Woche in der 6. Jahrgangsstufe (n=102) eingesetzt wurden.

In einer zweiten Studie wurde das Schreibflüssigkeitstraining mit jüngeren Schüler:innen erprobt. Dazu wurden die Aufgaben weiterentwickelt und für die Grundschule angepasst. Das Training wurde in einer Interventionsstudie mit Prä-Post-Kontrollgruppendesign mit 511 Dritt- und Viertklässler:innen

eingesetzt, anders als in der Vorgängerstudie wurde es von geschulten Lehrkräften der beteiligten Klassen mit Hilfe geskripteter Manuale durchgeführt.

Zur Messung der Schreibflüssigkeit wurde in beiden Studien im Prä- und Posttest die Anzahl der geschriebenen Buchstaben, Wörter und Sätze in einer bestimmten Zeit (Transkriptionsflüssigkeit) (Stephany et al. 2020) und mit digitalen Stiften die Anzahl der geschriebenen Wörter zwischen zwei mindestens zweisekündigen Pausen beim Schreiben eines Textes (Burstlänge, Formulierungsflüssigkeit) (Alves & Limpo, 2015) erhoben. Zusätzlich wurde u.a. die Textqualität erfasst. Die Implementationsgüte wurde in der zweiten Studie mithilfe von Unterrichtsbeobachtungen und Interviews mit den Lehrkräften untersucht. Positive Effekte auf die Schreibflüssigkeit zeigten sich sowohl in der 6. Jahrgangsstufe (F Gesamtwert Schreibflüssigkeit (1, 101)=36,4; p<.001) als auch in der 3. und 4. Klassenstufe (F Gesamtwert Schreibflüssigkeit (1, 509)=32,71; p<001). Trotz der insgesamt guten Gesamtergebnisse zeigten vertiefte Analysen jedoch große Unterschiede zwischen den teilnehmenden Schulen und zwischen einzelnen Schülergruppen: Einige Klassen und Kompetenzgruppen profitierten kaum oder gar nicht vom Training. Zwei Ursachenkomplexe sind hier nennenswert: (1) Individuelle Passung auf der Ebene der Schüler:innen. Die Aufgaben wirken differenzierend, das Training als Ganzes wird jedoch kaum differenziert durchgeführt. (2) Umsetzung und Passung auf der Ebene der Schulklasse bzw. Schule. Um die Ursachen der differenziellen Wirksamkeit des Trainings zu eruieren, wurden Voraussetzungen und Hintergrundvariablen einbezogen und die Ergebnisse des Trainings auf Teilstichprobenebene und für die Teilaspekte der Schreibflüssigkeit weiter analysiert. Zur Analyse der Umsetzung des Trainings wurden Beobachtungsdaten und Interviews zu Aussagen über die Implementationsgüte verdichtet.

Im Vortrag werden aus beiden Studien differenzielle Daten zu einzelnen Schülerkompetenzgruppen und Schulen gezeigt und vor dem Hintergrund zusammengefasster Skalen der Implementationsgüte diskutiert. Die Ergebnisse geben sowohl Hinweise für die Weiterentwicklung der eingesetzten Schreibflüssigkeitstrainings als auch allgemein für die Entwicklung und die Im-plementation von Trainings zur sprachlichen Förderung vor dem Hintergrund heterogener Lerngruppen und schulischer Rahmenbedingungen.

Von der Forschung in die Schule? Bedingungen eines erfolgreichen Transfers am Beispiel des Lesestrategietrainings FiLBY-3

Johannes Wild[1], Elisabeth Kraus[2], Sven Hilbert[2], Anita Schilcher[3]
[1]Lehrstuhl für Didaktik der deutschen Sprache und Literatur (Prof. Dr. Anita Schilcher), Universität Regensburg, [2]Methoden der empirischen Bildungsforschung, Universität Regensburg, [3]Lehrstuhl für Didaktik der deutschen Sprache und Literatur, Universität Regensburg

Theoretischer Hintergrund

International verfügen die deutschen Schülerinnen und Schüler in der Sekundarstufe über das höchste Lesestrategiewissen (Diedrich, Schiepe-Tiska, Ziernwald, Tupac-Yupanqui, Weis, McElvany & Reiss, 2019). Pearson & Cervetti (2017) bezeichnen die langfristige Integration von Strategien in den Schulalltag als Achillesferse von Lesestrategieprogrammen. Die Effektivität von Trainingsprogrammen hängt darüber hinaus stark von den Durchführungsbedingungen (Cheung & Slavin, 2016) und der Implementationstreue der Versuchsleitenden (Souvignier & Phillip, 2016) ab. Das Regensburger Lesestrategietraining FiLBY-3 (Wild, Schilcher & Steinert, 2020) setzt prototypisch

das im angloamerikanischen Raum bereits etablierte sechsstufige Vorgehen des Self-Regulated Strategy Developments (SRSD; Lienemann & Ried, 2006) um. Nach der Einführung von drei kognitiven Lesestrategien trainieren die Kinder zunächst mit Unterstützung der Lehrperson über mindestens vier Wochen hinweg mit Hilfe von knapp 70 gleich langen und gleich schweren Sachtexten.

In der vorliegenden Studie wird untersucht, unter welchen Bedingungen das Training die größte Wirksamkeit entfaltet. Dazu werden sowohl Daten einer Gruppe während des pandemiebedingten Distanzunterrichts als auch einer Gruppe im Präsenzunterricht danach untersucht.

Fragestellung

Es wird untersucht, unter welchen konkreten Bedingungen das Training in den beiden verschiedenen Settings wirksam ist:

- Wie adaptierten Grundschullehrkräfte das Training während der pandemiebedingten Schulschließungen im Frühjahr 2020 und wie manualgetreu führten sie es danach durch?

- Welche Parameter haben den größten (positiven) Einfluss auf die Entwicklung des Leseverstehens der Schülerinnen und Schüler?

Methode

An der longitudinalen Evaluation der FiLBY-Studie 2018-2020 nahmen insgesamt ca. 8.500 bayerische Schülerinnen und Schüler teil. Aufgrund der unvermittelten Schulschließungen war es vielen Lehrkräften in der dritten Jahrgangsstufe nicht mehr möglich, die Testungen durchzuführen. Ein großer Teil der Lehrkräfte der Experimentalgruppe konnte außerdem die FiLBY-3-Materialien nicht wie geplant einsetzen. Die Experimentalgruppe wurde daher für die Analysen geteilt: Die Schülerinnen und Schüler, die auch im Distanzunterricht weiterhin mit FiLBY arbeiteten, werden als Trainingsgruppe (TG) bezeichnet (n = 836). Diejenigen Kinder, die seit Beginn der Schulschließungen nicht mehr mit den FiLBY-Materialien arbeiteten, wurden zu einer Pseudokontrollgruppe (PG) zusammengefasst (n = 695). An einer Replikationsstudie nach den Schulschließungen 2021/22 nehmen insgesamt rund 1 300 Kinder teil. Zur Beantwortung der ersten Forschungsfrage werden deskriptive Statistiken erstellt. Für die zweite Forschungsfrage werden zwei hierarchisch lineare Regressionsmodelle geschätzt.

Ergebnisse und ihre Bedeutung

Die deskriptiven Analysen zeigen deutlich, wie die teilnehmenden Lehrerinnen und Lehrer das FiLBY-3-Training ihren Bedürfnissen während des Distanzunterrichts anpassten: Die Lehrkräfte ersetzten beispielsweise das zentrale Trainingselement des Modellierens durch ein Training ohne Strategien oder setzten auf Erklärtexte, Erklärvideos sowie die Fortführung von Leseflüssigkeitsübungen aus der zweiten Klasse. Auch nach den Schulschließungen zeigen sich deutliche Unterschiede bezüglich der Durchführungsbedingungen in den Klassen, etwa bei der Anzahl der gelesenen Texte.

Ein hierarchisches lineares Modell zeigt, dass sich die Trainingsanpassungen der Lehrkräfte während der Phase der Schulschließungen mehrheitlich ungünstig auf den Fortschritt der Kinder im Leseverstehen auswirken. Während die Kinder während des Trainings bei einer Strategieeinführung

durch Modellieren im Mittel 5.74 Punkte Fortschritt erzielen, zeigen die Anpassungen der Lehrkräfte signifikante Nachteile: Als besonders nachteilig erweist sich dabei ein (erneutes) Leseflüssigkeitstraining (-7.8 Punkte; $p < .001$), das Lesen ohne Strategien (-1.76 Punkte; $p < .001$) sowie Erklärvideos (0.31 Punkte; $p < .05$). Erklärtexte weisen hingegen keinen signifikanten Nachteil auf (1.34 Punkte; $p = .06$). Offensichtlich gelingt es den Kindern im Distanzunterricht nicht, sich Lesestrategien selbstständig anzueignen bzw. diese ohne Hilfe zu üben. Das Expertenmodell der Lehrkraft sowie deren Unterstützung in der Erwerbsphase scheinen demnach zentral für den Erfolg eines Lesestrategietrainings zu sein. Während des Präsenzunterrichts zeigt sich ebenfalls ein statistischer Vorteil der Experimentalgruppe. Ein weiteres Regressionsmodell zeigt, welche Parameter dabei den größten Einfluss auf den Leistungsfortschritt der Schülerinnen und Schüler haben. Daraus können Implikationen für den Erfolg von Trainingsprogrammen in der Praxis abgeleitet werden.

Förderung selbstregulierten Schreibens in Haupt- und Werkrealschulen

Seda Yilmaz Wörfel[1], Simone Jambor-Fahlen[1], Nora Fröhlich[2], Alexandra Dehmel[2], Benjamin Fauth[2], Michael Becker-Mrotzek[1]

[1]Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache, Universität zu Köln, [2]Institut für Bildungsanalysen Baden-Württemberg

Gute Schreibkompetenz ist für schulischen Erfolg, aber auch für die Teilnahme am gesellschaftlichen Leben unerlässlich. Das Erreichen der Mindeststandards im Schreiben (und Lesen) dokumentieren seit 2011 regelmäßig die Daten des IQB-Bildungstrends für die Jahrgänge 4 und 9: Von den 2016 getesteten ungefähr 30.000 Viertklässlerinnen und Viertklässlern aus über 1.500 Schulen erfüllten 53,9 % die Anforderungen der KMK-Regelstandards (Kompetenzstufe III) für den orthographischen Teilbereich, für 22,1 % konnten nur Leistungen unter der Kompetenzstufe II vermittelt werden, sie verfehlten damit die Mindesterwartungen (vgl. Weirich, Wittig & Stanat 2017, S. 133). Die Ergebnisse verdeutlichen die Notwendigkeit einer gezielten Förderung. Die Vermittlung selbstregulierter Schreibstrategien stellt einen der effektivsten Ansätze dar, die Schreibkompetenz von Schülerinnen und Schülern zu fördern und ihre Textproduktion zu unterstützen (Graham et al., 2012; Harris & Graham, 2017). Vor diesem Hintergrund zielt das Projekt „Die Textprofis“ darauf ab, die basalen Lese- und Schreibkompetenzen von Fünftklässlerinnen und Fünftklässlern an Haupt- und Werkrealschulen in Baden-Württemberg zu fördern. Das erstellte Fördermaterial konzentriert sich bei der Strategievermittlung im Bereich Schreiben auf das materialgestützte Schreiben von Zusammenfassungen. Diese Textsorte wurde gewählt, da sich das schriftliche Zusammenfassen bei der Förderung der Textqualität als äußerst effektiv erwiesen hat (Philipp, 2015). Die Schülerinnen und Schüler arbeiten außerdem mit einem digitalen Audiostift, der sie beim Lesen unterstützen soll. Das Förderkonzept umfasst flankierende Unterstützungsmaßnahmen in Form von Fortbildungen für die Lehrkräfte und Begleitung der Schulen durch Fachberaterinnen und Fachberater. Die Fachberaterinnen und Fachberater Unterricht wurden zuvor in einer Blended-Learning-Reihe zu Multiplikatoren weitergebildet. Das Förderprogramm wurde im Schuljahr 2021/22 an einer ausgewählten Stichprobe erprobt und im Rahmen einer Evaluationsstudie wissenschaftlich begleitet. In den folgenden Schuljahren soll ein Scaling-Up erfolgen.

In diesem Beitrag wird ein Teil des Konzepts mit Schwerpunkt auf der Förderung von Schreibstrategien sowie Evaluationsergebnisse zur Umsetzung des Förderprogramms in der Praxis zu folgender übergeordneten Forschungsfrage vorgestellt:

Was sind Erfolgsfaktoren und Herausforderungen bei der Umsetzung des Schreibstrategietrainings?

Das Training wird in einem quasi-experimentellen Design untersucht. Die Interventionsgruppe bestand aus 100 Lehrkräften aus 48 Schulen, denen eine Fortbildung zur Förderung der Schreibstrategien angeboten wurde. Die Vergleichsgruppe bestand aus 25 Lehrkräften aus 16 Schulen ohne vergleichbare Weiterbildung. In der Interventionsgruppe nahmen 1189 Schülerinnen und Schüler an dem Training teil (Vergleichsgruppe ohne Training N = ca. 340). 17 Fachberaterinnen und Fachberater Unterricht sowie 15 Fachberaterinnen und Fachberater Schulentwicklung begleiteten die Interventionsgruppe. Das Wissen der Schülerinnen und Schüler über Schreibstrategien wurde mittels eines Prä- und eines Posttests ermittelt (Knips et al., 2022). Zur Identifikation von Gelingensbedingungen bei der Umsetzung wurden qualitative und quantitative Daten erfasst. Nach dem Schreibstrategietraining wurden Schülerinnen und Schüler, Lehrkräfte sowie Fachberaterinnen und -berater beispielsweise mit Fragebögen zu verschiedenen Messzeitpunkten zur Einstellung gegenüber dem Training, zu den Unterstützungsmaßnahmen und zur Umsetzung befragt. Um Einblicke in den Umsetzungsprozess zu gewinnen, fanden regelmäßige Diskussions- und Feedbackrunden statt. Außerdem wurden Interviewdaten von Lehrkräften und Fachberaterinnen und -beratern gesammelt.

Erste deskriptive Ergebnisse zeigen, dass die Lehrkräfte sehr zufrieden mit der Fortbildung zur Förderung der Schreibstrategien waren. Etwas negativer, aber ebenfalls eher positiv bewertet wurde die Umsetzbarkeit in der Praxis sowie die Nützlichkeit der Fortbildung. Noch ausstehende Ergebnisse weiterer Analysen werden im Konferenzbeitrag berichtet.

Die erhobenen Daten sollen sowohl für die Wissenschaft als auch für die Praxis einen unmittelbaren Nutzen haben. Die gewonnenen (Zwischen)-Ergebnisse werden an die jeweiligen Akteure zurückgemeldet und dienen somit der datengestützten Anpassung des Förderprogramms (z.B. durch zusätzliche Unterstützungsmaßnahmen oder die Weiterentwicklung der Fortbildungsangebote in weiteren Durchläufen). Insgesamt sollen die Ergebnisse einen Beitrag zur Implementations- und Transferforschung leisten, insbesondere in Hinblick auf die Implementation wissenschaftsbasierter Fördermaßnahmen.

C-16: Simulationsbasiertes Lernen an der Hochschule – neue Erkenntnisse aus experimentellen und metaanalytischen Studien

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* R11 T00 D03

Simulationsbasiertes Lernen an der Hochschule – neue Erkenntnisse aus experimentellen und metaanalytischen Studien

Chair(s): **Frank Fischer** (Department Psychologie, Ludwig-Maximilians-Universität, München)
*Diskutant*in(nen):* **Cornelia Gräsel** (School of Education, Bergische Universität Wuppertal)

Simulationen gewinnen in vielen Bereichen der Hochschulbildung Bedeutung als Lernumgebungen, einschließlich der Mediziner- und Lehrkräftebildung. Dieses Symposium fokussiert Simulationen, die Aspekte professioneller Praxis repräsentieren und Studierenden durch die Bearbeitung authentischer Anforderungssituationen eine Annäherung an die Praxis ermöglichen (z.B. Grossman et al., 2009). Die Wirksamkeit von Simulationen beim hochschulischen Lernen ist mittlerweile gut belegt (z. B. Chernikova et al., 2020), wenngleich das Ausmaß des Lernerfolgs davon abhängt, inwieweit die Simulation bezüglich der Aufgabengestaltung im Sinne des repräsentationalen Scaffolding an die Lernvoraussetzungen adaptiert werden, (z. B. Cook et al., 2013) bzw. inwieweit zusätzliche Hilfestellungen im Lernprozess verfügbar sind (Chernikova et al., 2020).

Eine diesbezüglich viel beachtete Forschungsfrage ist, inwieweit und durch wen, Simulationen an die Lernvoraussetzungen angepasst werden sollten.

Der erste Beitrag des Symposiums untersucht wichtige Grundlagen für eine solche Anpassung. Betrachtet wird die Rolle zusätzlicher Reflexionsphasen als Scaffolding in simulationsbasierten Lernumgebungen für den Erwerb von kollaborativen Diagnosefähigkeiten. Der Beitrag prüft die Hypothese, dass Lernende mit weniger Vorwissen in geringerem Maß von Reflexionsphasen profitieren als Lernende mit mehr Vorwissen (Chernikova et al., 2020).

Im metaanalytischen zweiten Beitrag werden Mechanismen der Adaptation von Simulationen analysiert. Dabei steht die Frage im Mittelpunkt, wer die Anpassung vornimmt und bei welchen Lernvoraussetzungen die eigenständige Adaptierung durch die Lernenden selbst lernförderlich ist (vgl. Plass & Pawar, 2020).

Der dritte Beitrag betrachtet die Spezifität bzw. Sensitivität von Diagnosen im Kontext von Diagnosefähigkeiten Mathematiklehramtsstudierender und beleuchtet die in der Forschung zum simulationsbasierten Lernen vernachlässigte längsschnittliche Perspektive, verbunden mit der innovativen Frage, wie wiederholte Teilnahme und individuelles Vorwissen einer Entwicklung verschiedener Aspekte der Diagnoseakkuratheit zuträglich sind.

Mit Blick auf die mediale Gestaltung stehen derzeit vorrangig Fragen zur Verwendung von Video, AR und VR im Forschungsinteresse. Hier zeigen sich im Mittel positive Effekte für solche Technologien, die über reine Bildschirmpräsentationen der Aufgabe und Interaktionsmöglichkeiten mit Maus und Tastatur hinausgehen (Chernikova et al., 2020).

Im dritten Beitrag werden Effekte unterschiedlicher Präsentationsformate (interaktive Videosimulation vs. Rollenspielsimulation) von Simulationen untersucht. Der vierte Beitrag untersucht im Kontext der Interventionskompetenz im Inhaltsbereich Reanimation in der Medizinerausbildung, inwieweit Videos

aus der first-person-Perspektive zu Veränderungen im Debriefing und zu verbessertem Lernerfolg beitragen können.

Das Symposium umfasst Beiträge aus der Mediziner- und der Lehrkräftebildung. In der Diskussion soll u.a. die Frage erörtert werden, inwieweit Scaffolding (z.B. Reflexionsphasen, videogestütztes Debriefing) auch auf andere Fächer übertragen werden kann bzw. inwieweit die Generalisierbarkeit eher von der Art der zu erlernenden Fähigkeiten (z.B. Diagnostizieren, Intervenieren) als vom Fach abhängt.

Beiträge des Symposiums

Förderung kollaborativer Diagnosefähigkeiten für Lernende mit unterschiedlichem Vorwissen: Effekte von strukturierter Reflexion und Kollaborationsskripts in einer agentenbasierten medizinischen Simulation

Constanze Richters[1], Matthias Stadler[1], Anika Radkowitsch[2], Ralf Schmidmaier[3], Martin R. Fischer[4], Frank Fischer[5]
[1]Department Psychologie, Ludwig-Maximilians-Universität, München, [2]Didaktik der Mathematik, Leibniz-Institut für die Pädagogik der Naturwissenschaften und der Mathematik, Kiel, [3]Medizinische Klinik und Poliklinik IV, LMU Klinikum, Ludwig-Maximilians-Universität, München, [4]Institut für Didaktik und Ausbildungsforschung in der Medizin, LMU Klinikum, Ludwig-Maximilians-Universität, München, [5]Department Psychologie, Ludwig-Maximilians-Universität, München

Theoretischer Hintergrund

Simulationen mit zusätzlicher instruktionaler Unterstützung bieten eine Möglichkeit zur Förderung komplexer Fähigkeiten, wie z. B. kollaborative Diagnosefähigkeiten in der Medizin. Kollaborative Diagnosefähigkeiten umfassen für die Diagnosequalität entscheidende kognitive Aktivitäten wie das Generieren von Evidenzen und Hypothesen, das Ziehen von Schlussfolgerungen (Fischer et al., 2014) sowie das Elizitieren und Teilen von Evidenzen und Hypothesen (Radkowitsch et al., 2022). In der Medizin haben sich strukturierte Reflexion über frühe Verdachtsdiagnosen zur Verbesserung der Diagnosequalität (Mamede et al., 2014) und Kollaborationsskripts zur Förderung kollaborativer Diagnosefähigkeiten (Radkowitsch et al., 2021) als wirksam erwiesen. Darüber hinaus schlagen Vogel et al. (2017) vor, dass Kollaborationsskripts besonders wirksam sind, wenn sie mit inhaltsspezifischer Unterstützung ergänzt werden. Aufgrund dieser Befundlage wird erwartet, dass eine Kombination beider Unterstützungsformen vielversprechend ist. Die Wirksamkeit instruktionaler Unterstützung in Simulationen kann allerdings abhängig von Lernvoraussetzungen (z. B. Vorwissen) stark variieren. Neueste metaanalytische Befunde legen nahe, dass verschiedene Strukturierungsgrade von Unterstützungen mögliche Erklärungen für Wirkunterschiede sein können (Chernikova et al., 2020), da diese verschiedene Levels an Selbstregulation erfordern.

Fragestellung

In der vorliegenden Studie wurde untersucht, ob und inwiefern strukturierte Reflexion und Kollaborationsskripts den Erwerb kollaborativer Diagnosefähigkeiten in einer medizinischen Simulation unter Berücksichtigung des Vorwissens der Lernenden verbessern können.

Methode

151 Medizinstudierende (Nweiblich = 109) ab dem 5. Studienjahr wurden zufällig einer von vier Experimentalbedingungen (strukturierte Reflexion, Kollaborationsskript, beide Unterstützungsformen oder keine Unterstützung) zugewiesen. Die Studierenden bearbeiteten konsekutiv fünf Patient:innenfälle (Testfall, drei Lernfälle mit Lernunterstützung, Testfall) mithilfe einer

agentenbasierten Radiologin. Zu Beginn jedes Falles erhielten alle Studierenden eine Patient:innenakte mit relevanten klinischen Informationen. Teilnehmende der Reflexionsbedingung wurden anschließend dazu aufgefordert, Symptome und Befunde aus der Patient:innenakte mit Verdachtsdiagnosen zu verknüpfen. Im nächsten Schritt stellten alle Studierenden eine radiologische Untersuchungsanforderung, indem sie der Radiologin zur Begründung Patient:inneninformationen und Verdachtsdiagnosen mitteilen konnten. Sofern die Untersuchungsanforderung ausreichend begründet wurde, teilte die Radiologin ihre Untersuchungsergebnisse. Bei nicht ausreichender Begründung erhielten Teilnehmende der Kollaborationsskriptbedingung Aufforderungen mit Metawissen darüber, welche Informationen für die Radiologin relevant sind. Alle Studierenden schlossen den Fall mit Abgabe und Begründung einer finalen Diagnose ab. Wir berechneten die Qualität diagnostischer Austauschprozesse anhand von Evidenzen Teilen (Anteil geteilter relevanter Patient:inneninformationen an allen relevanten) und Hypothesen Teilen (Anteil geteilter relevanter Verdachtsdiagnosen an allen geteilten) als abhängige Variablen. Darüber hinaus erfassten wir das Vorwissen der Studierenden anhand von Single-Choice- und Key-Feature- (Page et al., 1995) Fragen sowie textbasierten Patient:innenfällen aus der Inneren Medizin bzw. Radiologie. Die Analyse erfolgte mittels multipler Regressionsmodelle.

Ergebnisse

Strukturierte Reflexion und Vorwissen zeigten keine signifikante Interaktion auf Evidenzen Teilen, jedoch einen signifikanten Effekt auf Hypothesen Teilen (Lernende mit hohem Vorwissen profitierten von strukturierter Reflexion, Lernende mit geringem Vorwissen nicht). Kollaborationsskripts und Vorwissen zeigten keine signifikanten Interaktionseffekte auf die untersuchten diagnostischen Austauschprozesse. Strukturierte Reflexion, Kollaborationsskripts und Vorwissen zeigten einen signifikanten Interaktionseffekt auf Evidenzen Teilen (Lernende mit hohem Vorwissen profitierten von einer Kombination strukturierter Reflexion und Kollaborationsskripts, Lernende mit geringem Vorwissen von Kollaborationsskripts), jedoch keinen signifikanten Effekt auf Hypothesen Teilen.

Diskussion

Strukturierte Reflexion und Kollaborationsskripts tragen dazu bei, Austauschprozesse in Simulationen zu verbessern. Lernende mit hohem Vorwissen profitieren von strukturierter Reflexion mit und ohne Kollaborationsskripts (Matthäus-Effekt; Walberg & Tsai, 1983; Chernikova et al., 2020), während Lernende mit geringem Vorwissen ausschließlich von Kollaborationsskripts profitieren (Radkowitsch et al., 2021). Diese Befundmuster deuten darauf hin, dass eine gering strukturierte Reflexion für Lernende mit hohem Vorwissen geeignet ist, während stärker strukturierte Kollaborationsskripts für Lernende mit niedrigem Vorwissen geeignet sind. Vorliegende Studie stützt damit die Annahme, dass die Wirksamkeit instruktionaler Unterstützung in Simulationen für Lernende mit unterschiedlichem Vorwissen aufgrund verschiedener Strukturierungsgrade variiert (Chernikova et al., 2020). Damit wird ein Beitrag geleistet zur aktuellen Frage, inwieweit die Anpassung instruktionaler Unterstützung in Simulationen an individuelle Lernvoraussetzungen die Wirksamkeit simulationsbasierten Lernens erhöht.

Nicole Heitzmann[1], Olga Chernikova[1], Matthias Stadler[1], Doris Holzberger[2], Tina Seidel[2], Frank Fischer[1]
[1]Department Psychologie, Ludwig-Maximilians-Universität, München, [2]TUM School of Social Sciences and Technology, Technische Universität München

Theoretischer Hintergrund

Die Wirksamkeit des Lernens mit Simulationen wurde in verschiedenen Bereichen der Hochschulbildung empirisch gezeigt (z. B. Chernikova et al., 2020, Theelen et al., 2019). Ein weniger einheitliches Bild ergibt sich, wenn zusätzliche Unterstützungsmaßnahmen betrachtet werden. So waren die Wirksamkeit einzelner Unterstützungsmaßnahmen wie beispielsweise zusätzlicher Reflexionsphasen in Abhängigkeit vom Vorwissen der Lernenden wirksam (Chernikova et al., 2020). Die vorliegende Metaanalyse zielt darauf ab, den Mehrwert der Adaptivität von instruktionaler Unterstützung in Simulationen für Lernende mit geringem und hohem Vorwissen beim Erwerb von komplexen Fähigkeiten zu ermitteln. Sie stützt sich auf Theorien und Befunde zum Wissens- und Fertigkeitserwerb (z. B. Van Lehn, 1996) sowie zum Scaffolding für unterschiedliche Phasen des Wissens- und Fertigkeitserwerbs (z. B. Hmelo-Silver et al., 2006). Die Kategorisierung der verschiedenen Arten bzw. Adaptivitätsstrategien der instruktionalen Unterstützung in Simulationen basiert auf dem von Plass & Pawar (2020) vorgeschlagenen Rahmenmodell zur Adaptivität von Lernumgebungen.

Fragestellungen

(FF1) Wie werden Adaptivitätsstrategien in Simulationen in der Hochschulbildung umgesetzt?

(FF2) Inwieweit können Adaptivitätsstrategien Lernergebnisse bei Lernenden mit höherem und niedrigerem Vorwissen in Simulationen in der Hochschulbildung fördern?

Methode

Die Metaanalyse basiert auf einem aktualisierten Datensatz der Studie von Chernikova et al. (2020). Dabei wurde dieselbe Suchstrategie angewandt und Artikel bis Dezember 2020 eingeschlossen. Die Metaanalyse basiert auf 400 Effekten aus 146 veröffentlichten Studien. Ein Random Effects-Modell sowie Hedges g wurden verwendet (Borenstein et al, 2009). Die folgenden Adaptivitätsstrategien wurden kodiert: (1) Fixierte Adaptierung basierend auf dem Pre-test, (2) Adaptierung basierend auf der Performanz in der Simulation, (3) Adaptation der Zeit und des Umfangs der Unterstützung durch die Lernenden, (4) Adaptation der Reihenfolge durch die Lernenden, (5) Lernende können ihr Verhalten während der Bearbeitung z.B. nach Feedback anpassen, (6) Die Bearbeitung der Simulation kann wiederholt werden, bis das Lernziel erreicht ist.

Ergebnisse

Der Haupteffekt von Simulationen auf den Erwerb von Kompetenzen war g=.99, SE=.07. Dieser wurde als Referenzpunkt für die Einschätzung der Effektivität der einzelne Adaptivitätsstrategien genutzt. (1) Fixierte Adaptierung wurde in keiner der Primärstudien genutzt (N=0). (2) Performanzbasierte Adaptierung (N=29) erhöhte tendenziell die Effekte hauptsächlich für Lernende mit hohem Vorwissen (genutzt: g=1.31, SE=.30; nicht genutzt: g=0.92, SE=.09). (3) Adaptation der Unterstützung durch die Lernenden (N=18) hatte tendenziell positive Effekte für Lernende mit hohem (Lernende: g=1.16, SE=.36; Lehrende/System: g=.98, SE=.10) und negative Effekte für Lernende mit geringem Vorwissen (Lernende: g=.43, SE=.12; Lehrende/System: g=.77, SE=.14). (4) Adaptation der

Reihenfolge durch die Lernenden (N=108) war eine tendenziell effektive Strategie (genutzt: g=1.09, SE=.15; nicht genutzt: g=.83, SE=.08). (5) Verhaltensanpassung der Lernenden (N=10) war tendenziell eine Strategie mit negativen Effekten (genutzt: g=.84, SE=.30; nicht genutzt: g=.92, SE=.07). (6) Wiederholung der Simulation (N=100) war für Lernende mit geringem Vorwissen eine sinnvolle Strategie (genutzt: g=.90, SE=.21; nicht genutzt g=.66, SE=.12).

Diskussion

Zusammenfassend lässt sich sagen, dass die Effekte von Simulationen durch Adaptierung der instruktionalen Unterstützung verstärkt werden können. Nicht jede Strategie wird dabei gleichermaßen genutzt. Einige Strategien zeigen je nach Lernvoraussetzungen spezifische Effekte. So zeigte sich für Lernende mit geringem Vorwissen, dass Sie von Strategien, die ein hohes Maß an Selbststeuerung bzw. Metakognition benötigen, nicht profitieren. Für diese Lernenden scheinen Adaptierungsstrategien sinnvoller zu sein, durch die sie mehr üben können (z.B. Wiederholung der Simulation). Bei Lernenden mit hohem Vorwissen scheint die performanzbasierte Adaptierung durch die Simulationen zu gelingen, so dass sie passendere oder mehrere Aufgaben zur Verfügung haben. Bei Lernenden mit geringem Vorwissen scheint dies in Simulationen weniger gut zu gelingen. Limitationen dieser Metaanalyse sind (1) eine relativ grobe Kategorisierung der Adaptivitätsstrategien und (2) eine knappe Beschreibung der Simulationen in den Primärstudien. Insgesamt deutet diese Metaanalyse darauf hin, dass Adaptivität in Simulationen vor allem dann positive Effekte erzielt, wenn sich die Adaptationsstrategie am Vorwissen der Lernenden orientiert.

Entwicklung diagnostischer Kompetenz zur Dezimalbruchrechnung: Akkuratheit, Sensitivität und Spezifität von Diagnosen in einer simulationsbasierten Lernumgebung

Stephanie Kron[1], Daniel Sommerhoff[2], Stefan Ufer[1]
[1]Lehrstuhl für Didaktik der Mathematik, Ludwig-Maximilians-Universität, München, [2]Didaktik der Mathematik, Leibniz-Institut für die Pädagogik der Naturwissenschaften und der Mathematik, Kiel

Forschungsstand

Der Erwerb diagnostischer Kompetenzen ist Teil der Lehrkräftebildung (Kultusministerkonferenz, 2022). Um diesen Kompetenzerwerb möglichst realitätsnah zu gestalten, werden authentische Lernumgebungen empfohlen, in denen Studierende die praktische Anwendung professionellen Wissens aktiv einüben können (Grossman et al., 2009). In der Lehrkräftebildung werden hierfür insbesondere Videosimulationen (van Es & Sherin, 2008), im Medizinstudium traditionell häufig Rollenspiele genutzt (Lane & Rollnick, 2007). Rollenspiele zeichnen sich durch authentische Handlungsmöglichkeiten aus, wohingegen Videosimulationen eine stärker vorstrukturierte Problemsituation bieten. In Simulationen erfasste Performanzmaße können dazu dienen, die Kompetenzentwicklung zu analysieren. Im Bereich diagnostischer Kompetenzen wird hierzu insbesondere die Akkuratheit der Diagnose berücksichtigt, also ihre Übereinstimmung mit einer Referenzdiagnose (Südkamp et al., 2012). Für eine differenzierte Erfassung sind dabei zwei Aspekte zu berücksichtigen: Lehrkräfte müssen sowohl Stärken als auch Schwächen der Schüler:innen diagnostizieren können. Lernstandsdiagnosen sollten daher Fähigkeitsbereiche zuverlässig erfassen (sog. Spezifität), sowie Fehlvorstellungen und Problembereiche korrekt identifizieren (sog. Sensitivität). Differenzierte längsschnittliche empirische Untersuchungen von Lernstandsdiagnosen, die über die reine Akkuratheit auch deren Spezifität und Sensitivität berücksichtigen, fehlen bisher.

Fragestellung

Folgende Fragen werden fokussiert: 1. Welchen Einfluss haben fachliches und fachdidaktisches Vorwissen auf Akkuratheit, Spezifität und Sensitivität diagnostischer Urteile? 2. Wie verändern sich diese Leistungsfacetten bei wiederholter Teilnahme in simulierten Diagnoseinterviews? 3. Welchen Einfluss hat die Durchführung der Simulation als Rollenspiel bzw. als Videosimulation auf diese Leistungen und deren Entwicklung?

Methode

Mathematiklehramtsstudierende nahmen in der Lehrkraftrolle an vier simulierten Diagnoseinterviews teil, wobei durchgehend ein festes von zwei parallelen Präsentationsformaten (Videosimulation, N=66; Rollenspielsimulation, N=53) randomisiert zugewiesen wurde (N=476 Simulationen)(Marczynski et al., 2022). Im simulierten Diagnoseinterview sollten diagnostische Informationen gesammelt und eine Diagnose über den Lernstand einer Schülerin/eines Schülers erstellt werden. Die Schülerrolle wurde im Rollenspielformat von geschulten Projektmitarbeiter:innen gespielt. Für die Videosimulation wurden geskriptete Videos von Sechstklässler:innen erstellt. Zur Messwiederholung wurden vier verschiedene Schülerprofile genutzt und randomisiert zugewiesen.

Nach jedem Diagnoseinterview bewerteten die Studierenden den Lernstand der Schülerin/des Schülers hinsichtlich neun Einzelfacetten in einem geschlossenen Antwortformat (Antwortmöglichkeiten: „sicher beherrscht“, „nicht sicher beherrscht“, „keine Diagnose möglich“). Antworten der Kategorie „keine Diagnose möglich“ wurden als Indikator für fehlende Sicherheit im Treffen diagnostischer Urteile interpretiert und in den weiteren Analysen ausgeschlossen. Für die Akkuratheit der Diagnose wurden alle Bewertungen berücksichtigt, die für das jeweilige Schülerprofil korrekt abgegeben wurden. Für die Spezifität wurden ausschließlich die Einzelfacetten herangezogen, die das jeweilige Schülerprofil beherrscht, für die Sensitivität diejenigen, die das Schülerprofil nicht beherrscht. Vor der ersten Simulation wurde das Fachwissen und das fachdidaktische Wissen erhoben. Die Auswertung erfolgte auf Ebene der Einzelfacetten mittels verallgemeinerter linearer Mischmodelle.

Ergebnisse

Für die Akkuratheit der Diagnose zeigten sich positive Zusammenhänge zum Fachwissen ($B=0.13;p<.01$) und fachdidaktischen Wissen ($B=0.12;p<.05$). Effekte der wiederholten Teilnahme zeigten sich ab der dritten Simulation ($B_{(1.vs3.Sim)}=0.42;p<.0001;B(1.vs4.Sim)=0.46;p<.0001$). Dasselbe Befundmuster ergab sich für die Sensitivität. Für die Spezifität konnten keine signifikanten Effekte des Vorwissens beobachtet werden, jedoch eine positive Entwicklung ab der zweiten Simulation ($B(1.vs2.Sim)=0.46;p<.05;B(1.vs3.Sim)=0.45;p<.05;B(1.vs4.Sim)=0.56;p<.01$). Für die Sicherheit im Treffen diagnostischer Urteile zeigten sich keine signifikanten Zusammenhänge zum Vorwissen, jedoch ebenfalls Effekte der wiederholten Teilnahme ($B(1.vs2.Sim)=0.45;p<.01;B(1.vs3.Sim)=0.77;p<.0001;B(1.vs4.Sim)=0.97;p<.0001$). Signifikante Effekte des Präsentationsformats ergaben sich nicht.

Diskussion

Dass die Sicherheit ab der zweiten Simulation zunahm, lässt vermuten, dass die Lernenden zunehmend in ihre Diagnosen vertrauen. Die Ergebnisse deuten darauf hin, dass insbesondere Akkuratheit und Sensitivität mit dem Vorwissen variieren, während Spezifität und Sicherheit keine derartigen Zusammenhänge zeigten. Offen bleibt damit, wieso mehr Vorwissen nicht dazu beiträgt, auch Fähigkeiten zu identifizieren. Alle vier Maße waren für Fortschritte über die vier Simulationen hinweg sensitiv, und es konnte eine positive Entwicklung beobachtet werden, die sich allerdings nicht nach Präsentationsformat unterschied. Trotz der gegenüber der Videosimulation erweiterten Handlungsmöglichkeiten scheint das Rollenspiel also keine unterschiedliche Kompetenzentwicklung anzuregen.

Förderung von Reanimationsskills durch simulationsbasiertes Lernen: Welche Vorteile bringt der Einsatz von first-person-view-Videos im Debriefing?

Martin Gartmeier, Nina Soellner, Matthias Eiberle, Rainer Haseneder, Dominik Hinzmann, Christian Schulz, Stefan Rath, Pascal Berberat
Klinikum rechts der Isar, Technische Universität München

Hintergrund und Fragestellung

Simulationsbasiertes Lernen hat hohen Stellenwert hinsichtlich der Vorbereitung junger Ärzt*innen auf die klinische Praxis. Dies gilt besonders für Akutsituationen in denen ärztliches Handeln über Leben und Tod entscheiden kann – z.B. durch kardiopulmonale Reanimation (CPR). Solche Situationen stellen hohe Anforderungen an die Handelnden, zumal wenn diese im Team agieren. In CPR-Simulationen sind Debriefing und Feedback daher kritische Faktoren für Kompetenzentwicklung. Aktuell werden vielversprechende technologiebasierte Feedbackformen erprobt, u.a. unter Nutzung von first-person-view-Videos (FPV). Diese werden mithilfe einer Kamerabrille aufgenommen und erfassen so die subjektive Perspektive einer handelnden Person (Wolf et al., 2019). Feedback, das Lernende nach einer absolvierten CPR-Simulation bekommen kann dazu beitragen, entscheidende Momente während der Reanimation zu erkennen und kritische Reflexion bzgl. der eigenen Performanz anzuregen. Da FPV-Videos die individuelle Lernendenperspektive einfangen gehen wir entsprechend dem Konzept des Lernprozess-Scaffolding (Fischer et al., in press) davon aus, dass Lernende dadurch besonders gut in ihrer Kompetenzentwicklung unterstützt werden. Ziel einer randomisiert-kontrollierten Prä-Post-Studie mit Medizinstudierenden war daher die Beantwortung der Frage, ob und in welcher Form Feedback mit/ohne FPV-Videos die Lernwirksamkeit eines simulierten Reanimationsszenarios beeinflusst (Soellner et al., 2022). Im Fokus der Studie standen dabei Outcomes auf objektiver und klinischer Ebene (siehe Methode).

Methode

Die Studie stand im Rahmen eines notfallmedizinischen Lehrmoduls an dem 108 Medizinstudierende im dritten klinischen Studienjahr teilnahmen. Die Studierenden absolvierten ein simuliertes Reanimationsszenario (Prätest) mit Resusci Anne Advanced SkillTrainer-Simulationspuppen (Laerdal Medical), sie wurden randomisiert einer von zwei Debriefing-Bedingungen zugewiesen (Dozierendengeleitetes Feedback mit/ohne FPV-Videos). Nach dem Debriefing absolvierten die Studierenden ein zweites simuliertes Reanimationsszenario (Posttest) mit vergleichbarer Aufgabenstellung. Die Teilnehmenden in beiden Treatment-Bedingungen trugen Kamerabrillen, zusätzlich wurden die Szenarien mit Überblickskameras videographiert.

Objektive Outcomes der Studie waren Situationsbewusstsein und Teamarbeit als wichtige nicht-technische Aspekte der Reanimationsperformanz, sie wurden von zwei notfallmedizinisch erfahrenen Ratern auf Basis der Videoaufzeichnungen mit Überblickskameras mit hoher Reliabilität eingeschätzt. Situationsbewusstsein umfasste die Aspekte Informationssammlung, Erkennen und Verstehen und Voraussicht (ICC=.72); Teamarbeit umfasste Teamkommunikation, Teamkooperation, positive Einstellung und Flexibilität (ICC=.92). Auf der Ebene klinischer Outcomes nutzten wir die Effective-Compression-Ratio (ECR; Greif et al., 2013) als Maß für die Reanimationsqualität. Dieser Index wird aus verschiedenen über die Simulationspuppe erfassten technischen Parametern gebildet (u.a. Kompressionstiefe bei der Reanimation, Kompressionsrate und Handposition). Zudem erfassten wir die zeitliche Dauer je Reanimationsteam bis relevante Reanimationsereignisse ausgeführt wurden (erste Thoraxkompression und erste Defibrillation).

Ergebnisse

In der Studie zeigten sich statistisch bedeutsame Trainingseffekte, aber keine Unterschiede zwischen Interventions- und Kontrollgruppe. Zwischen Prä- und Posttest-Szenario zeigten sich signifikante Effekte bzgl. beider objektiver Outcomes, der Teamarbeit ($p<0.001$; $F=13.99$) und des Situationsbewusstseins ($p<0.001$; $F=11.4$). Auf klinischer Ebene verbesserte sich die ECR in beiden Treatmentgruppen, der Unterschied war aber nur in der Kontrollgruppe signifikant ($p<0.005$). Die Zeit bis zum Starten der Thoraxkompression verkürzte sich in Kontroll- und Interventionsgruppe statistisch signifikant ($p<0.005$ bzw. $p< 0, 001$). Die Dauer bis zur ersten Defibrillation, welcher eine sehr hohe Priorität in den einschlägigen Leitlinien zugeordnet wird, verkürzte sich ebenfalls signifikant in beiden Trainingsgruppen (Kontrolle: $p<0.01$; Intervention: $p<0.001$). Der Einfluss der Gruppenzugehörigkeit (Kontrolle / Intervention) war bei keiner der Variablen statistisch signifikant.

Diskussion

Die Studie zeigt, dass sich FPV-Videos als moderne Debriefingtechnologie mit vertretbarem Aufwand in einen curricularen notfallmedizinischen Simulationskurs integrieren lassen. Trotz genereller Kompetenzgewinne konnten wir jedoch keinen zusätzlichen Vorteil von FPV-Videos gegenüber dozierendengeleiteten Debriefings zeigen. Zusätzlich erhobene Freitextantworten deuten darauf hin, dass FPV-Videos Vorteile für den subjektiven Erkenntnisgewinn im Bereich nicht-technischer Fähigkeiten mit sich zu bringen (z.B. Teamarbeit). Um diese Vorteile jedoch in bessere klinische Performanz im Bereich CPR umzusetzen, gehen wir davon aus, dass eine intensivere Begleitung FPV-basierter Debriefings durch Dozierende notwendig ist, z.B. um relevante Ereignisse in den Videos zu identifizieren und um deren Bedeutung zu reflektieren.

C-17: Der Umgang von Lehrkräften mit Herausforderungen in der Bildung für nachhaltige Entwicklung (BNE)

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* T03 R04 D10

Der Umgang von Lehrkräften mit Herausforderungen in der Bildung für nachhaltige Entwicklung (BNE)

Chair(s): **Marit List** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany), **Johannes Hartig** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany)
*Diskutant*in(nen):* **Marco Rieckmann** (Universität Vechta)

Durch Bildung für nachhaltige Entwicklung (BNE) sollen Kompetenzen von Schüler*innen gefördert werden, die nachhaltiges gesellschaftliches Handeln ermöglichen (Rieckmann, 2018). Zwar ist BNE in den Lehrplänen deutscher Schulen bereits verankert, jedoch bisher kaum in den Unterricht integriert (Rieckmann & Holz, 2017). Ein Grund mag sein, dass Bildung für nachhaltige Entwicklung (BNE) an Lehrkräfte neue Herausforderungen stellt, für die die Lehramtsausbildung derzeit zumindest noch nicht in vollem Umfang vorbereitet (Rieckmann & Holz, 2017).

Herausforderungen in der BNE betreffen unter anderem den Umgang mit den eigenen Einstellungen und Überzeugungen in Bezug auf globale Probleme im Kontext von (nicht-)nachhaltiger Entwicklung (Grund & Brock, 2020) und die Notwendigkeit zur Anwendung neuer Unterrichtskonzepte (Grundmann, 2017). Auch die Motivation, sich neue Themengebiete und Unterrichtsmethoden anzueignen, sowie der Umgang mit strukturellen Hindernissen stellen zentrale Herausforderungen bei der Umsetzung von BNE dar (Grundmann, 2017).

Das Symposium beschäftigt sich damit, wie Lehrkräfte diese Herausforderungen wahrnehmen und wie sie mit diesen in ihrem schulischen Handeln innerhalb und außerhalb des Unterrichts umgehen. Unser Symposium nähert sich dem Thema aus unterschiedlichen fachwissenschaftlichen Disziplinen (Fachdidaktik, Pädagogik, Psychologie) und unter Nutzung unterschiedlicher Forschungsdesigns (mixed methods, qualitativ). Die vorgestellten Beiträge zeigen verschiedende Aspekte der Bedeutung von Lehrkräften und unterrichtlichem Handeln für gelingende Unterrichtsgestaltung im Sinne der BNE auf: Der erste Beitrag untersucht förderliche und hinderliche Faktoren bei der Umsetzung von BNE-Praktiken im Schulalltag und im Unterricht. Die Autor*innen stellen Ergebnisse einer qualitativen Interviewstudie mit Lehrkräften zu deren Erfahrungen mit der Umsetzung von BNE vor. Der zweite Beitrag befasst sich mit unterrichtlichen Handeln. Die Autor*innen stellen die Ergebnisse einer qualitativen Studie mit Lehrkräften zum didaktischen Umgang mit Unsicherheiten im Themenfeld Klimawandel im Geographieunterricht vor. Der dritte Beitrag befasst sich ebenfalls mit unterrichtlichem Handeln: Basierend auf einer qualitativen Typologie untersuchen die Autor*innen anhand von Daten aus einer internationalen Bildungsvergleichsstudie die Unterrichtspraxis von Lehrkräften bei der Bearbeitung globaler Themenkomplexe im internationalen Vergleich.

Die Ergebnisse der in den Beiträgen vorgestellten Forschung zeigen verschiedene Herausforderungen auf, die bei der Integration von BNE in das Schulgeschehen berücksichtigt werden müssen und möglichst auch in der Lehramtsausbildung adressiert werden sollten, um bereits hier auf veränderte Anforderungen an das Unterrichtshandeln einzugehen. Die BNE-spezifischen Befunde zum Umgang von Lehrkräften mit Herausforderungen sind auch auf andere Themengebiete übertragbar, und so können die vorgestellten Ergebnisse auch außerhalb der BNE-Forschung

relevant sein. Ein abschließender Diskussionsbeitrag adressiert, welche Limitationen die vorgestellten Arbeiten haben und wo noch weiterer Forschungsbedarf in diesem Themengebiet besteht.

Beiträge des Symposiums

Umsetzungsmöglichkeiten für Bildung für nachhaltige Entwicklung (BNE) aus Lehrkraftperspektive: eine qualitative Interviewstudie

Marit List, Johannes Hartig
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany

Theoretischer Hintergrund: Bildung für nachhaltige Entwicklung (BNE) ist an Schulen eine wichtige Querschnittsaufgabe, und Lehrkräften kommt bei der Umsetzung eine Schlüsselrolle zu (Gräsel et al., 2013). BNE soll nicht nur in der Unterrichtsgestaltung, sondern auch im Schulleben berücksichtigt werden, etwa durch nachhaltige Gestaltung des Schulalltags oder durch Projekte außerhalb des regulären Unterrichts (Valentin, 2021). Damit werden durch BNE vielfältige neue Anforderungen an Lehrkräfte gestellt (Grundmann, 2017).

Die BNE-bezogenen Tätigkeiten von Lehrkräften können sowohl als Form von Unterrichts- und allgemeinen schulischen Handeln aber auch als Form von nachhaltigkeitsfördernden Verhalten betrachtet werden. Aus beiden Perspektiven lässt sich die Annahme ableiten, dass BNE-Tätigkeiten von Lehrkräften, innerhalb und außerhalb des Unterrichts, durch die persönlichen Einstellungen und Überzeugungen, aber auch durch wahrgenommene Verhaltenskontrolle und Handlungsbarrieren beeinflusst sind (Ajzen, 1991; Baumert & Kunter, 2006; Gifford, 2011). Handlungsbarrieren finden sich im Schulkontext zum Beispiel in den räumlichen Bedingungen (Ausstattung), der strukturellen (Schulcurriculum) wie auch der sozialen Schulorganisation (Zusammenarbeit im Kollegium). Bisherige Studien bestätigen, dass neben den Einstellungen und Überzeugungen Handlungsbarrieren zentral für die Umsetzung sind (Buddeberg, 2014; Grundmann, 2017). Die hier vorgestellte Studie knüpft an diese Forschung an und untersucht die für Lehrkräfte zentralen Faktoren von gelingender BNE-Umsetzung.

Fragestellung: Diese Studie geht der Frage nach, welche Faktoren für eine erfolgreiche Umsetzung von BNE aus Lehrkraftperspektive relevant sind. Dazu wurde die individuelle Perspektive der Lehrkräfte betrachtet – im Hinblick auf die eigenen Kompetenzen und Handlungsmöglichkeiten sowie aus hinderliche und förderliche Aspekte in der Schulorganisation. Es wurden angehende und im Schuldienst tätige Lehrkräfte verschiedener Schultypen und außerschulische BNE-Akteur*innen interviewt und nach ihren Erfahrungen mit der Umsetzung von BNE befragt. Ziel der Studie ist es, über Schulformen und Akteursgruppen hinweg zentrale individuelle und strukturelle Faktoren, die bei der Umsetzung von BNE in der Schule eine Rolle spielen, zu explorieren.

Methode: Datengrundlage bilden 16 Interviews mit Lehrkräften (n = 11), Personen, die in der außerschulischen BNE-Arbeit tätig sind (n = 3) und Lehramtsstudierende mit einem besonderen Studienfokus BNE (n = 2). Die Lehrkräfte haben unterschiedliche Fachhintergründe (u.a. Biologie, Deutsch, Mathematik) und sind an verschiedenen Schultypen (Grundschule, Realschule, Gymnasium, berufsbildende Schule) tätig. Die Interviews waren teilstrukturiert und folgten einem Leitfaden mit Fragen zu den persönlichen BNE-Erfahrungen im Unterricht, den eigenen nachhaltigkeitsbezogenen Einstellungen und Überzeugungen, sowie weiterer Faktoren, die eine Umsetzung von BNE

begünstigen oder verhindern. Die Frageimpulse waren offen gestaltet und die einzelnen Themen wurden in ihrer Ausgestaltung an den Interviewverlauf angepasst. Die transkribierten Interviews wurden inhaltsanalytisch ausgewertet (Kuckartz, 2016). Die Kategorien wurden sowohl deduktiv auf Grundlage des aktuellen Forschungsstands (Buddeberg, 2014; Grundmann, 2017) als auch induktiv anhand der spezifischen Inhalte der Interviews entwickelt.

Ergebnisse und ihre Bedeutung: Die Interviews zeigen, dass BNE prinzipiell in allen Fächern, Schulformen und Altersstufen aufgegriffen werden kann. Zentral für die Integration von BNE in den Unterricht sind das Interesse der Lehrkraft an Nachhaltigkeitsthemen und ein hohes Engagement bei der Umsetzung. Darüber hinaus ist ein engagiertes Kollegium von hoher Relevanz. Wesentliche Hinderungsgründe bei der Umsetzung sind fehlende Materialien, fehlendes Wissen und fehlende Zeit für Materialerstellung. Selten genannt werden hingegen fehlende finanzielle Mittel. Die Interviews zeigen eine große Varianz in den eigenen Erfahrungen mit der Umsetzung von BNE, dies betrifft sowohl die behandelten Themen, die Verortung von BNE innerhalb oder außerhalb des Unterrichts wie auch das wahrgenommene Interesse der Schüler*innen.

Die Ergebnisse der Studie zeigen, dass für eine erfolgreiche Gestaltung von Unterricht und Schulalltag im Sinne der BNE sowohl individuelle Faktoren auf Seiten der Lehrkräfte und Schüler*innen, aber auch strukturelle Faktoren (Schulorganisation) und die Zusammenarbeit im Kollegium relevant sind. In der Forschung zu schulischer BNE und bei Maßnahmen zu ihrer Förderung sollten demzufolge sowohl die Ebene der individuellen Lehrkraft als auch der schulische Kontext Berücksichtigung finden.

Der professionelle Umgang mit handlungs- und entscheidungstheoretischen Unsicherheiten von Geographielehrer*innen im Kontext des Klimawandels

Melissa Hanke, Angelika Paseka, Sandra Sprenger
Universität Hamburg

Theoretischer Hintergrund: Angesichts der zunehmenden Komplexität von gesellschaftlichen Veränderungen sowie der darin eingebetteten Ambiguität und Flüchtigkeit von Wissen hat Unsicherheit zugenommen und steht immer mehr im Fokus von Bildung und Wissenschaft (Unger, 2019), unter anderem auch im Bereich der Erziehungswissenschaften (Böing, 2016). In der Geographiedidaktik wird Unsicherheit sowohl auf theoretischer als auch auf unterrichtspraktischer Ebene thematisiert, vor allem bei zukunftsbezogenen Themen aus dem Bereich Bildung für Nachhaltige Entwicklung (BNE). Im Rahmen einer BNE ist der Klimawandel ein zentrales Inhaltsfeld, welches im Rahmen dieser Studie als exemplarisches Unterrichtsthema dient (Condeza-Marmentini & Flores-González, 2019; Perkins et al., 2018).

Bei der Thematisierung des Klimawandels im Unterricht kann zwischen verschiedenen Formen von auftretenden Unsicherheiten differenziert werden: zum einen treten wissens- und wissenschaftstheoretische Unsicherheiten auf (Busch & Osborne, 2014; IPCC, 2022; Ruggeri, 2011) und zum anderen treten handlungs- und entscheidungstheoretische Unsicherheiten bei Schüler*innen und Geographielehrer*innen auf (Melville & Pilot, 2014; Sterling, 2010; Tauritz, 2019). Letztere stehen im Fokus dieses Artikels.

Handlungs- und entscheidungstheoretische Unsicherheiten sind im Geographieunterricht in jeder Unterrichtssituation auf individueller und interaktioneller Ebene konstitutiv, z. B. dadurch, dass die Lehrer*innen nicht wissen, welche Gedanken bei den Schüler*innen während des Unterrichts auftreten, welche Lernprozesse ablaufen oder wie die Schüler*innen auf die Handlungen der Lehrer*innen reagieren (Floden & Clark, 1988; Melville & Pilot, 2014; Paseka et al., 2018). Diese Unsicherheiten könnten sich insbesondere bei kontrovers diskutierten, komplexen und von Unsicherheit geprägten Unterrichtsthemen wie dem Klimawandel verstärken und eine Herausforderung für Geographielehrer*innen darstellen (Chang, 2014; Chang & Wi, 2018).

Aus professionstheoretischer Perspektive müssen Geographielehrer*innen mit diesen handlungs- und entscheidungstheoretischen Unsicherheiten in ihrer Unterrichtsplanung und -durchführung umgehen. Dabei ist eine große Bandbreite an möglichen Umgangsweisen denkbar, die differente Auswirkungen auf den Unterricht haben können, beispielweise ist eine Marginalisierung des Unterrichtsthemas Klimawandel denkbar, ebenso aber auch eine aktive Nutzung von Unsicherheiten (Bähr et al., 2019). Im Umgang mit diesen Unsicherheiten ist, beruhend auf der Wissenssoziologie nach Mannheim, nicht das explizite Wissen der Geographielehrer*innen handlungsleitend, sondern die impliziten, atheoretischen Orientierungen (Nohl, 2017), die im Zentrum dieses Forschungsprojektes stehen.

Fragestellung: Basierend auf dem dargelegten theoretischen Hintergrund wurde der folgenden übergeordneten Forschungsfrage nachgegangen: Woran orientieren sich Geographielehrer*innen beim Umgang mit handlungs- und entscheidungstheoretischen Unsicherheiten im Kontext des Klimawandels?

Um dieser Frage nachzugehen, wurden zunächst die Unsicherheitsbereiche und -momente identifiziert, anschließend die Orientierungen der Geographielehrer*innen rekonstruiert und mögliche Absorptionsmechanismen analysiert.

Methode: Um die Orientierung von Geographielehrer*innen zu rekonstruieren, wurde ein qualitativ-rekonstruktives Untersuchungsdesign gewählt. Es wurden narrative Einzelinterviews mit 15 Geographielehrer*innen an Hamburger Gymnasien geführt, die sich hinsichtlich ihrer Berufserfahrung unterscheiden. Aufgrund der Corona-Pandemie wurden die Interviews digital mit Hilfe von Zoom-Videokonferenzen geführt. Die Interviews wurden mit Hilfe der dokumentarischen Methode ausgewertet, um im Rahmen der formulierenden und reflektierenden Interpretation die Orientierungsrahmen und -schemata zu konstruieren. Aktuell erfolgt die Typenbildung (Nohl, 2017).

Ergebnisse und ihre Bedeutung: Die Typenbildung weist eine große Bandbreite von der Unsicherheitsaversion bis zur Nutzung von Unsicherheit aus. Sie wird aktuell abgeschlossen und im Rahmen des Vortrags vorgestellt.

Lehrkräfte als Akteur*innen globaler Transformation im Kontext einer Bildung für nachhaltige Entwicklung

Dorothea Taube[1], **Jana Costa**[2]
[1]jana.costa@lifbi.de, [2]Leibniz-Institut für Bildungsverläufe

Theoretischer Hintergrund: Als ‚Change Agents' kommt Lehrkräften in der Vorbereitung junger Menschen auf die Herausforderungen einer globalisierten Welt eine besondere Bedeutung zu (Merki und Werner, 2011; OECD und Asia Society, 2018). Wie Lehrkräfte globale Lerngegenstände

verhandeln, kann den Umgang zukünftiger Generationen mit globalen Entwicklungen, Problemstellungen und damit auch globale Transformationsprozesse beeinflussen. Aktuelle empirische Befunde verweisen in diesem Zusammenhang darauf, dass die Vorstellungen der Lehrkräfte über den Unterrichtsgegenstand und die Art der didaktischen Umsetzung in vielfältiger Weise miteinander verbunden sind. Dies wird auch in der qualitativen Studie zum Umgang von Lehrkräften mit globalisierungsbezogenen Themen- und Fragestellungen deutlich (Taube, 2022). In der Untersuchung wurden die Erzählungen und Beschreibungen von 17 Lehrkräften zu ihrer Handlungspraxis hinsichtlich der Frage nach wiederkehrenden Formen der Bearbeitung globaler Themenstellungen analysiert. Ziel der Arbeit war es, Muster zu identifizieren und idealtypisch zu beschreiben, die sich bei Lehrkräften in der Beschäftigung mit Themen aus globaler Perspektivität zeigen.

Methodisches Vorgehen: Die aus den empirischen Daten entwickelte sinngenetische Typologie und die generierten Hypothesen (Taube, 2022) dienen in diesem Beitrag als theoretischer Ausgangspunkt für eine systematische Reanalyse der Daten der International Civic and Citizenship Education Study (ICCS). Diese internationale Vergleichsstudie untersucht den Beitrag von Bildungssystemen mit Blick auf demokratische (Entwicklungs-) Prozesse und fokussiert Wissen, Überzeugungen und Handlungsperspektiven von Schülerinnen und Schüler im Kontext politischer Bildung. Die darin thematisierten Inhalte und Aspekte sind anschlussfähig an den Bereich einer Bildung für nachhaltigen Entwicklung. Überdies bezieht die Studie als Kontextvariablen Lehrkräftebefragungen mit ein (Schulz et al., 2017).

Erkenntnisinteresse: Im Fokus des Beitrags stehen wünschenswerte Aspekte lehrkräftebezogener Handlungspraxis (Taube, 2022). Dazu werden (1) die Zielorientierungen (mit Blick auf die Aspekte der Reflexivität, des Wissens und der Partizipation) und (2) die Unterrichtsmethoden von Lehrkräften im internationalen Vergleich betrachtet, die im Horizont der qualitativen Typologie als erstrebenswert erscheinen. In den Blick kommt darüber hinaus (3) das Zusammenspiel zwischen den spezifischen Zielorientierungen und dem Einsatz von bestimmten Unterrichtsmethoden. Die Ergebnisse auf Ebene der Lehrkräfte werden (4) vor dem Hintergrund der Befunde der Schüler*innen (mit Blick auf die Lücke zwischen Wissen und Handeln) näher betrachtet.

Ergebnisse: Bezüglich der Zielorientierungen der Lehrkräfte deutet sich an, dass die erstrebenswerte Kombination aus Reflexivität, Wissen und Handeln in allen Ländern den niedrigsten Anteil ausmacht. Es lässt sich allerdings über alle Länder hinweg eine positive Korrelation zwischen den Zielorientierungen Wissen & Reflexivität beobachten, d.h. wenn Lehrkräfte Wissen als zentral anerkennen, dann geben sie auch häufiger Reflexivität als bedeutsam an. Die Muster der Kombination der Zielorientierungen unterscheiden sich dabei zwischen den Ländern. Im Horizont der abgefragten Unterrichtsmethoden zeigt sich, dass solche Unterrichtsmethoden, die offene und partizipative Settings eröffnen (z.B. Projektarbeit außerhalb der Schule), Perspektivenvielfalt zulassen (z. B. Rollenspiele) und die Schüler*innen in den Mittelpunkt stellen (z.B. Vorschläge von Unterrichtsthemen von Schüler*innen) verhältnismäßig selten eingesetzt werden. Diese und weitere Ergebnisse der Reanalyse werden im Beitrag vorgestellt und in Hinblick auf die Bedeutung des Handelns von Lehrkräften im Kontext einer Bildung für nachhaltigen Entwicklung diskutiert.

C-18: Stereotype im Lehrer*innenzimmer und ihre Auswirkungen auf das heterogene Klassenzimmer

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* T03 R02 D39

Stereotype im Lehrer*innenzimmer und ihre Auswirkungen auf das heterogene Klassenzimmer

Chair(s): **Sabine Glock** (Bergische Universität Wuppertal, Deutschland), **Hannah Kleen** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main)
*Diskutant*in(nen):* **Ursula Kessels** (Freie Universität Berlin)

Die Bildungsbenachteiligung von Schüler*innen mit Migrationshintergrund und aus niedrigen sozialen Schichten ist nicht nur für Deutschland immer noch ein Thema (Autor*innengruppe Bildungsberichterstattung, 2022). Dabei werden in der Forschung einerseits niedrigere Schulleistungen dieser Schüler*innen und andererseits die Rolle der Lehrkräfte diskutiert (Glock et al., 2020; Olczyk et al., 2016). Forschung zeigt, dass Stereotype und Einstellungen von Lehrkräften gerade in Bezug auf den Migrationshintergrund auf das Lehrkraftverhalten wirken (z.B. Glock & Böhmer, 2018). Im ersten Beitrag des Symposiums berichten die Autor*innen von einer Studie, in der angehende Lehrkräfte die Ursachen für gute sowie schlechte Schülerleistungen in Abhängigkeit des Migrationshintergrundes des Schülers angaben. Angehende Lehrkräfte berichten sowohl für den leistungsstarken Schüler als auch für den Schüler ohne Migrationshintergrund mehr Ursachen. Für die Leistungen schwacher Schüler und Schüler mit Migrationshintergrund scheint es entsprechend weniger Erklärungsbedarf zu geben. Möglicherweise stellen Stereotype diese Attributionen bereit. Im zweiten Beitrag des geplanten Symposiums können die Autorinnen feststellen, dass Erwachsene mit Migrationshintergrund vermehrt versuchen, Schulabschlüsse über den zweiten Bildungsweg nachzuholen, aber auch, dass diese häufig die Ausbildung abbrechen. Die längsschnittliche Untersuchung zeigt, dass gerade für Erwachsene mit Migrationshintergrund die Lehrkraft-Schüler*innen-Beziehung im Laufe der Ausbildung schlechter wird, was mit den Abbruchsquoten in Verbindung gebracht werden kann. Stereotype können sich negativ auf die Lehrkraft-Schüler*innen-Beziehung auswirken (Babad, 2007). Die negativen Auswirkungen von Stereotypen resultieren in den letzten Jahren vermehrt in der Frage nach den individuellen Merkmalen von Lehrkräften, die diese aushebeln könnten (z.B. Kleen et al., 2019). Der dritte Beitrag widmet sich zweier solcher Merkmale, dem eigenen Migrationshintergrund von Lehrkräften sowie deren Selbstwirksamkeitserwartungen und Überzeugungen zum Unterrichten heterogener Lerngruppen. Die Autorinnen berichten davon, dass Lehrkräfte mit eigenem Migrationshintergrund selbstwirksamer und auch enthusiastischer für das Unterrichten in heterogenen Lerngruppen sind. Unabhängig vom eigenen Migrationshintergrund wird zudem deutlich, dass diese hohen Selbstwirksamkeitserwartungen auf den warmen Umgang mit leistungsschwachen Schüler*innen wirken. So können Stereotype in den Hintergrund gedrängt werden. Jedoch ist es schwierig, Lehrkräfte mit diversem Hintergrund in den Schuldienst zu bekommen (Statistisches Bundesamt, 2010). Dies liegt nicht nur an dem geringen Zugang zur Hochschule, sondern auch in den Zugangsvorrausetzungen, ein Lehramtsstudium zu beginnen. Der letzte Beitrag aus Luxemburg zeigt hier einen starken Verbesserungsbedarf, weil die Hürden für Personen mit Migrationshintergrund über die starke Betonung der Sprachkompetenzen teilweise zu hoch sind. Diese Studie zeigt, dass die Kompetenz in der Amtssprache stärker im Vordergrund steht als pädagogische Kompetenzen. Damit scheinen sich Stereotype im heterogenen Klassenzimmer weiter fortzusetzen, da ein heterogenes Lehrer*innenzimmer noch in weiter Zukunft ist.

Beiträge des Symposiums

„Da hättest du dich aber mehr anstrengen können!“- Kausalattributionen von Lehramtsstudierenden zur Leistung von Schüler*innen mit Migrationshintergrund

Meike Bonefeld[1], Hannah Kleen[2], Tobias Baumann[3], Sabine Glock[3]
[1]Universität Konstanz, [2]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main, [3]Bergische Universität Wuppertal

In vielen Bildungssystemen erzielen beispielsweise Schüler*innen mit Migrationshintergrund schlechtere schulische Leistungen, und folglich entwickeln Lehrkräfte niedrige Erwartungen an diese Schüler*innengruppe (Jussim & Harber, 2005). In diesem Zusammenhang können nicht nur die Urteile der Lehrkräfte über verschiedene Leistungsergebnisse (Bonefeld & Dickhäuser, 2018, Glock, 2016) sondern auch die Urteile darüber wo Ursachen für Leistung zu finden sind durch Stereotype beeinflusst sein. Solche Kausalattributionen können der Attributionstheorie von Weiner zufolge internal oder external sowie variabel oder stabil sein (Weiner, 2000) und sich für Erfolg und Misserfolg unterscheiden. Studien, die Kausalattributionen in den Blick nehmen zeigten, dass Lehrkräfte das Scheitern von Schüler*innen in erster Linie auf schülerinterne Faktoren, wie z. B. Anstrengung, zurück und nur selten auf external-variable Aspekte, die z.B. die Unterrichtsqualität betreffen (Wang & Hall, 2018). Es konnte überdies gezeigt werden, dass sich die Kausalattributionen von Lehrkräften abhängig vom Migrationshintergrund der Schüler*innen zu Ungunsten von Schüler*innen mit Migrationshintergrund unterscheiden. Lehrkräfte führen den akademischen Erfolg von Schüler*innen mit Migrationshintergrund stärker auf internal-variable Faktoren (z.B. Anstrengung) zurück als bei Schüler*innen ohne Migrationshintergrund und berücksichtigen externale Faktoren (z. B. ihre eigene Unterrichtspraxis) stärker, wenn sie die erfolgreichen Leistungen von Schüler*innen mit Migrationshintergrund bewerten (Wang & Hall, 2018). Auch Versagen wird eher auf internale, meist stabile Ursachen zurückgeführt und externale Aspekte in der Ursachenzuschreibung durch Lehrkräfte vernachlässigt (Reyna, 2008). Solch differentielle Kausalattributionen abhängig vom Migrationshintergrund können sich auch auf die Beurteilung der Leistungen verschiedener Schüler*innen durch die Lehrkräfte auswirken. Wird das Versagen von Schüler*innen mit Migrationshintergrund in erster Linie auf internale, stabile Faktoren zurückgeführt, wird die Kompetenz der Schüler*innen von Lehramtsstudierenden weniger positiv beurteilt (Glock & Kleen, 2021). In der vorliegenden Studie beurteilten 175 Lehramtsstudierende (143 weiblich, Mage24.64, SDage= 3.93) Schülervignetten, die entweder einen Schüler beschrieben, der entweder unterdurchschnittliche oder überdurchschnittliche Schulleistungen und schulisches Engagement zeigte und einen oder keinen Migrationshintergrund aufwies hinsichtlich der Ursachen für dessen Leistung. Eine 2x (Leistungsniveau: niedrig vs. hoch) x 2 (Migrationshintergrund: mit vs. ohne) x 4 (Kausalattributionen: internal-stabil vs. internal-variabel vs. external-stabil vs. external-variabel) ANOVA mit Messwiederholung für den letzten Faktor zeigte einen signifikanten Haupteffekt des Leistungsniveaus des Schülers, $F(1, 171) = 236,88$, $p < .001$, $\eta p2 = 0,58$. Im Allgemeinen gaben die Teilnehmer*innen mehr Kausalattributionen für die leistungsstarken als für die leistungsschwachen Schüler an. Die ANOVA ergab auch einen signifikanten Haupteffekt des Migrationshintergrundes, $F(1, 171) = 9,50$, $p = .002$, $\eta p2 = 0,06$. Bei Schülern ohne Migrationshintergrund stimmten die Lehramtsstudierenden mehr Attributionen zu als bei Schülern mit Migrationshintergrund. Die Interaktion zwischen

Leistungsniveau und Migrationshintergrund war nicht signifikant, F(1, 171) = 0,77, p = .38, ηp2 = 0,00. Der Haupteffekt der Kausalattributionen war signifikant, F(3, 513) = 63.68, p < .001, ηp2 = 0.27. Die Lehramtsstudierenden stimmten internal-stabilen Attributionen (M = 3.35, SD = 1.06) mehr zu als internal-variablen und external-stabilen. Sie stimmten auch internal-stabilen Attributionen mehr zu als external-variablen Gründen. Internal-variable Attributionen wurden mehr befürwortet als external-stabile, aber weniger als external-variable Gründe. External-stabile Attributionen fanden weniger Zustimmung als external-variable Attributionen. Die Interaktion zwischen Leistungsniveau und Kausalattributionen erreichte Signifikanz, F(3, 513) = 44.15, p < .001, ηp2 = 0.21. Bei jeder der Kausalattributionen stimmten die Lehramtsstudierenden den Aussagen für den leistungsstarken Schüler mehr zu als für den leistungsschwachen Schüler. Die Interaktion zwischen Migrationshintergrund und Kausalattributionen war ebenfalls signifikant, F(3, 513) = 3.49, p = .016, ηp2 = 0.02. Die Lehramtsstudierenden bestätigten mehr external-stabile Attributionen für die Schüler ohne Migrationshintergrund als für die Schüler mit Migrationshintergrund. Alle anderen Vergleiche waren nicht signifikant. Die dreifache Interaktion war nicht signifikant, F(3, 513) = 0,17, p = .92, ηp2 = 0,001. Die Studie zeigte somit interessante und unerwartete Ergebnisse, welche im Beitrag diskutiert werden.

Migrant*innen im zweiten Bildungsweg – Entwicklung von Lehrerbeziehungen und Bildungserfolg

Claudia Schuchart, Doris Bühler-Niederberger, Angelika Siebel
Bergische Universität Wuppertal

Theoretischer und empirischer Hintergrund: Eine steigende Zahl von Erwachsenen mit Migrationshintergrund wählt den zweiten Bildungsweg, um die Hochschulreife nachzuholen. Überdurchschnittlich viele von ihnen brechen diesen Bildungsweg jedoch wieder ab (Bellenberg 2019). Für eine Erklärung dieses Befundes kommen verschiedene Ansätze in Frage. In diesem Vortrag möchten wir uns auf die Lehrerbeziehung konzentrieren. Im Rahmen des Sozialkapitalansatzes wird angenommen, dass eine unterstützende und vertrauensvolle Beziehung zu Lehrkräften Schüler*innen dabei hilft, ihre Bildungsziele zu erreichen (Roth 2014), was auch für den zweiten Bildungsweg gezeigt werden konnte (Schuchart & Bühler-Niederberger 2019). Allerdings weisen andere Studien darauf hin, dass Lernende mit Migrationshintergrund stärkere Schwierigkeiten haben, eine positive Beziehung zu Lehrkräften aufzubauen (z.B. Holland, 2015, Schuchart & Bühler-Niederberger 2019). Unter anderem beeinträchtigen stereotypisierende Vorstellungen von Lehrkräften die Beziehung zu Lernenden mit Migrationshintergrund (Holland, 2015; Yeager 2021). Es fehlen bislang jedoch quantitative Studien, die dieses Phänomen mit Bezug zu Bildungszielen aus einer Entwicklungsperspektive untersuchen.

Fragestellung: Im vorliegenden Beitrag möchten wir untersuchen, wie sich unter erwachsenen Lernenden mit und ohne Migrationshintergrund die Beziehung zu Lehrkräften unter Berücksichtigung von Diskriminierungserfahrungen im Zeitverlauf entwickelt und inwieweit diese Entwicklung mit dem Abbruchrisiko im Zusammenhang steht.

Methode: Unser Datensatz umfasst 160 Lernende im Alter von ca. 24 Jahren im zweiten Bildungsweg, die im Abstand von zwei Jahren u.a. zu Aspekten der Qualität ihrer Beziehungen zu

Lehrkräften im zweiten Bildungsweg befragt worden sind. Der Migrationsstatus wird über die Frage erfasst, wo der Befragte und seine Eltern geboren sind und die Nähe zur Mehrheitskultur (Musik, Film, Traditionen; Eigenentwicklung, 3 bipolare Items, 7 Antwortkategorien, alpha = .60). Die Beziehung zu den Lehrkräften wird über 8 Items aus der Perspektive der Lernenden erfasst, die u.a. Aspekte der Fairness, Wertschätzung, Interesse, Erfolgserwartungen und Disziplinierung erfragen. Es werden Kontrollvariablen wie Alter und Geschlecht verwendet. Die Auswertung erfolgt unter Verwendung von Regressionen in folgenden Schritten: Zunächst wird die Entwicklung verschiedener Aspekte der Lehrerbeziehung a) in Abhängigkeit vom Migrationshintergrund und b) im Zusammenhang mit dem allgemeinen Abbruchrisiko regressionsanalytisch geprüft. Anschließend wird c) das Abbruchrisiko nach Migrationsstatus und Nähegrad zur Mehrheitskultur untersucht und es wird d) versucht, Differenzen durch die Aufnahme von Aspekten der Lehrerbeziehung zu erklären. Zusätzlich werden e) Interaktionseffekte zwischen Migrationserfahrung bzw. Nähegrad und Lehrerbeziehung auf das Abbruchrisiko geprüft.

Ergebnisse: Bisher vorliegende Ergebnisse zeigen, dass sich a) die Beziehung zwischen Lernenden mit Migrationshintergrund und Lehrkräften bei teilweise geringerem Ausgangsniveau weniger günstig entwickelt als für Lernende ohne Migrationshintergrund. Dies betrifft Aspekte von Wertschätzung, Fairness und Erfolgserwartungen, wobei der Nähegrad zur Mehrheitskultur eine größere Rolle spielt als der Migrationsstatus an sich. b) Eine Verschlechterung aller untersuchten Aspekte der Lehrerbeziehung steht mit einer Erhöhung des Abbruchrisikos im Zusammenhang. c) Ein höheres Abbruchrisiko steht eher mit dem Nähegrad zur Mehrheitskultur im Zusammenhang, weniger mit dem Migrationsstatus an sich. d) Die Aspekte der Lehrerbeziehung erklären die mit der Ferne zur Mehrheitskultur steigende Abbruchneigung. e) Insbesondere mangelnde wahrgenommene Erfolgserwartungen haben mit steigender Ferne zur Mehrheitskultur einen zunehmenden Effekt auf das Abbruchrisiko.

Diskussion: Die Ergebnisse legen nahe, dass im zweiten Bildungsweg zur Erlangung der Hochschulreife insbesondere der Nähegrad zur Mehreitskultur relevant für die Beziehungsentwicklung zu Lehrkräften und, darüber vermittelt, für das Abbruchrisiko ist. Dies verweist darauf, dass nicht so sehr askriptive Merkmale wie der Migrationshintergrund an sich, sondern Prozessmerkmale wie die hier erhobene mehrheitskulturelle Nähe von Bedeutung sind. Wie auch in Studien im Bereich der Sozialkapitalforschung hervorgehoben, ist die wahrgennommene Erfolgserwartung besonders für Lernende mit größerer mehrheitskultureller Ferne von Bedeutung für den Bildungserfolg. Obgleich Lehrkräfte im zweiten Bildungsweg insgesamt eine stärker lernerorientierte Haltung aufweisen (Schuchart et al. 2019), scheint es zumindest aus der Wahrnehmung der Lernenden Unterschiede zu geben, die auch in der Stereotypenforschung von der Lehrerseite her belegt sind.

Überzeugungen und Selbstwirksamkeitserwartungen im Unterrichten leistungsheterogener Lerngruppen. Unterschiede zwischen ethnischen Minoritäts- und Majoritätslehramtsstudierenden und Auswirkungen auf die Adaptivität des Lehrkraftverhaltens in Lehrkraft-Lernenden-Dyaden

Madeleine Kreutzmann, Karoline Koeppen, Theresa Mros, Bettina Hannover
Freie Universität Berlin

Heterogenität im Lehrer:innenzimmer kann sich einerseits darin zeigen, dass Lehrpersonen selbst Minoritätsgruppen – z.B. in Bezug auf ihren ethnischen Hintergrund – angehören, und andererseits in heterogenitätsbejahenden Einstellungen und Überzeugungen gegenüber den Lernenden. Lehrkräfte, die selbst einer Minorität angehören, können insbesondere für benachteiligte Schüler:innen mehr Chancengleichheit herstellen, weil sie signalisieren, dass sie die Diversität der Schüler:innen als eine positive Ressource betrachten und selbst gerne Lernende unterrichten wollen, die aus Minoritätsgruppen kommen. So fanden beispielsweise Schachner et al. (2018), dass ein solches positives Diversitätsklima vermittelt über das Gefühl von Zugehörigkeit (Belonging) sich günstig auf das akademische Selbstkonzept und die Leistungen von Schüler:innen auswirkte.

In Studie 1 berichten wir über die Befragung von 500 Master-Lehramtsstudierenden, die selbst einer ethnischen Minorität oder der ethnischen Majorität angehörten. Wie erwartet gaben Studierende ethnischer Minoritätsgruppen eine stärkere Bereitschaft, einen stärkeren Enthusiasmus, eine höhere subjektive Kompetenz und eine stärkere Selbstwirksamkeit, Klassen mit einem hohem Anteil von Schüler:innen mit Migrationshintergrund zu unterrichten, an.

Wenig ist darüber bekannt, ob und wenn ja, auf welche Weise sich positive heterogenitätsbezogene Einstellungen von Lehrkräften auf die Schüler:innen auswirken. Schotte et al. (2022) fanden beispielsweise keine Zusammenhänge zwischen entsprechenden Überzeugungen der Lehrkräfte und motivationalen und Leistungsoutcomes auf Seiten der Schüler:innen. In unserer zweiten Studie untersuchten wir die Hypothese, dass der vermittelnde Mechanismus die Adaptivität des Verhaltens der Lehrperson in Abhängigkeit der Kompetenz des Lernenden ist. Achtundachtzig Grundschul-Lehrkräfte beschrieben ihr Verhalten gegenüber jeweils fünf Kindern ihrer Klasse (insgesamt 440 Schüler:innen) auf den Dimensionen Communion (Wärme, Herzlichkeit) und Agency (Lenkung, Überwachung). Erwartungsgemäß zeigte sich, dass Lehrkräfte mit hoher Selbstwirksamkeit im Unterrichten heterogener Lerngruppen ihr Verhalten stärker in Abhängigkeit der Leistungen des jeweiligen Kindes anpassten als Lehrkräfte mit geringer heterogenitätsbezogener Selbstwirksamkeit. Lehrkräfte, die gegenüber dem Unterrichten in leistungsheterogenen Gruppen positiv eingestellt waren, zeigten erwartungsgemäß eine höhere Communion gegenüber den Schüler:innen als Lehrkräfte mit kritischer Einstellung, und zwar insbesondere gegenüber Kindern mit schwachen Leistungen.

Wir diskutieren die Ergebnisse in Hinblick auf die Frage, wie mehr Minoritätsstudierende für das Lehramt gewonnen und wie Lehrkräfte noch stärker darin unterstützt werden können, ihr Verhalten adaptiv an den Kompetenzen des jeweiligen Lernenden auszurichten.

Diversifikation im Klassenzimmer – ein Einwegspiegel?

Ineke Pit-ten Cate[1], Salvador Rivas[1], Gilbert Busana[2]
[1]Luxembourg Centre for Education Testing (LUCET), Faculty of Humanities, Education, and Social Sciences, University of Luxembourg, [2]Department of Education and Social Work, Faculty of Humanities, Education, and Social Sciences, University of Luxembourg

Über mehrere Jahrzehnte hinweg ist weltweit eine Zunahme der Diversität der Schüler*innen auf allen Bildungsstufen zu verzeichnen. Diese studentische Vielfalt spiegelt sich jedoch nur teilweise im Lehrkörper wider, in dem die Mehrheitsgruppen der Länder noch immer überrepräsentiert sind. Da Lehramtsstudiengänge hauptsächlich auf Universitätsebene angeboten werden, kann diese

Überrepräsentation der Mehrheitsgruppe Bildungsungleichheiten widerspiegeln, die zu niedrigeren schulischen Leistungsergebnissen für Schüler*innen aus Minderheitengruppen (OECD, 2019) und letztendlich zu ihrer Unterrepräsentation in höheren Ebenen der Sekundar- oder Hochschulbildung (Baumert & Schümer, 2002; Waters et al., 2013) führen. Obwohl die Unterrepräsentation bestimmter Schüler*innengruppen mit individuellen Unterschieden im akademischen Potenzial zusammenhängen kann, können solche soziale Ungleichheiten auch aus einem Mangel an positiven Rollenvorbildern für bestimmte Schüler*innengruppen und einem unzureichenden multikulturellen Bewusstsein in Bildungskontexten resultieren (vgl. Akbaba et al., 2013; Magaldi et al., 2018).

Diese Vorbildfunktion soll den Einfluss negativer, auf Stereotypen basierender Erwartungen der Lehrkräfte auf die schulischen Leistungen der Schüler*innen verringern, da die Lehrkraft mit Migrationshintergrund ein „lebender Beweis“ für Bildungserfolg ist (vgl. Marx & Roman, 2002). In dieser Hinsicht hat die Forschung gezeigt, dass Schüler*innen aus Minderheitengruppen oft besser abschneiden, wenn sie Lehrer*innen haben mit denen sie sich identifizieren können oder die deren soziokulturellen Hintergrund verstehen (Dee, 2005; Egalite et al., 2015; Hachfeld et al., 2012). Zudem ist zu erwarten, dass Lehrkräfte mit Migrationshintergrund Schüler*innen mit ähnlichem Hintergrund anders wahrnehmen und positiver bewerten als Lehrkräfte ohne Migrationshintergrund. Diese Vorstellung wird durch Erkenntnisse gestützt, die zeigen, dass sowohl Lehramtsstudierende als auch erfahrene Lehrkräfte mit Migrationshintergrund eine positivere Einstellung gegenüber Schülern mit Migrationshintergrund haben als Lehrkräfte aus der Mehrheitsgruppe (Glock & Kleen, 2019; Kleen et al., 2019).

Daher versuchen viele Länder seit Jahren, den Anteil von Lehrkräften mit Migrationshintergrund zu erhöhen (Ingersoll & May, 2011; Rotter, 2012). Studien haben jedoch gezeigt, dass Lehramtsstudiengänge immer noch durch eine Unterrepräsentation von Studierenden mit Migrationshintergrund gekennzeichnet sind (Döll & Knappik, 2015; Edelmann, 2013).

Fragestellung:

In Luxemburg spiegeln sich diese Trends wider. In Luxemburg sind Schüler*innenkohorten zunehmend durch sozioökonomische, kulturelle und sprachliche Diversifizierung gekennzeichnet. Diese Vielfalt spiegelt sich jedoch nicht im Lehrkörper wider.

Die aktuelle Studie befasst sich mit der Forschungsfrage inwiefern das Auswahlverfahren für die Zulassung zum Lehramtsstudium - in Luxemburg umfasst die Zulassung zum Programm eine wettbewerbliche Aufnahmeprüfung zum Nachweis mathematisch-naturwissenschaftlicher Kompetenzen, sowie der drei Amtssprachen des Landes: Luxemburgisch, Deutsch und Französisch - Bewerbern und Bewerberinnen aus Mehrheits-/Minderheitengruppen, Vor- oder Nachteile gewähren kann.

Methode:

In der aktuellen Studie untersuchten wir die Diversität der Kandidaten, die sich für den Lehramtsstudiengang angemeldet haben, in Bezug auf die Konstellation der zugelassenen Lehramtsstudierenden. Dazu haben wir soziodemografische Daten für fünf Kohorten (2015-2019)

verwendet. Die soziodemografischen Daten umfassten Alter, Geschlecht, Nationalität und Muttersprache der Kandidaten, sowie Bildungsabschluss und Beruf der Eltern.

Ergebnisse und ihre Bedeutung

Ergebnisse einer logistischen Regressionsanalyse zeigten, dass die soziodemografischen Merkmale zwischen (nicht-) erfolgreichen Bewerbern und Bewerberinnen sich unterschieden, Chi2 (5, N=989) = 127.52, p<.001. Genauer gesagt haben männliche Bewerber, Bewerber*innen mit luxemburgisch als Erstsprache und Bewerber*innen mit Eltern mit höherem Bildungsniveau eine signifikant höhere Chance zum Lehramtsstudium zugelassen zu werden. Obwohl die Kohorte der registrierten Bewerber*innen recht vielfältig ist, ist die Vielfalt der Kohorte der zugelassenen Bewerber*innen geringer, da die Mehrheit aus einheimischen (basierend auf der 1. Sprache) und besser gebildeten Familien stammt.

Diese Ergebnisse weisen darauf hin, dass die Verzögerung bei der Diversifizierung des Lehrkörpers teilweise darauf zurückzuführen ist, dass der Schwerpunkt auf der Beherrschung der Amtssprachen des Landes liegt, da dies für Bewerber und Bewerberinnen mit Migrationshintergrund Nachteile mit sich bringt (s.a. Döll & Knappik, 2015). Um die Diversität der Lehrerschaft zu erhöhen, ist es daher notwendig, nicht nur die Verringerung von Bildungsungleichheiten in Betracht zu ziehen, sondern auch die Diversifizierung der Einstiegswege in die Lehrer*innenbildung.

C-19: School adjustment and success of recently arrived refugee and immigrant children in Germany

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* T03 R03 D89

School adjustment and success of recently arrived refugee and immigrant children in Germany

Chair(s): **Francesca Ialuna** (Universität Duisburg-Essen), **Sandy Chwastek** (Ruhr-Universität Bochum, Fakultät für Psychologie, Child and Family Research), **Sebastian Vogel** (TU Dortmund, Institut für Schulentwicklungsforschung)
*Diskutant*in(nen):* **Linda Juang** (Universität Potsdam)

In the last decades, a large number of families with young children migrated to or asked for asylum in Germany (BMF, 2017; 2019). Hence, the adjustment of recently arrived refugee children as well as of children of immigrant origin is a critical matter for schools and societies. In Germany, especially the elementary school is a crucial phase, as it prepares for the transition into the tracked secondary school system and might thus shape children's future academic trajectories (e.g., Schindler, 2017). Several migration-related experiences (e.g., exposure to war, discrimination, resettlement in a new society) can however challenge refugee and immigrant children's academic development (Bronstein & Montgomery, 2011). Therefore, it is crucial to examine how children who are recently migrated to Germany adjust at school and which factors may play a role in enhancing their adjustment.

This symposium aims at highlighting different aspects of the school adjustment and success of recently arrived refugee and immigrant children. All contributions focus on the particular situation of refugee and immigrant pupils in the last year of elementary school in predominantly low-income, multi-ethnic neighborhoods in Germany. Taking an ecological approach (Bronfenbrenner, 1979; Suárez-Orozco et al., 2018), we discuss the perspectives of parents, teachers, peers as well as the children themselves to gain a clearer insight on how these pupils are and adjust in school.

The first presentation addresses differences in the secondary school recommendations at the end of primary school between children of different migration generations. Additionally, further aspects of school success of recently arrived children are regarded, while also considering other family background variables as well as individual and family level mediators.

The second contribution highlights the importance of the family context for children's school adjustment. In particular, it explores the relations between parental resources, values, home involvement, school involvement of refugee and immigrant parents, and their children's school adjustment.

In the third presentation, we examine the quality of children's relationships with teachers and peers. Specifically, the study addresses the protective role of these relationships for refugee and immigrant children's school adjustment in spite of the acculturative challenge of learning the language of instruction.

In light of the findings, implications for educational interventions and future research will be discussed.

Beiträge des Symposiums

School success of refugee and immigrant children at the end of primary school

Sebastian Vogel[1], Justine Stang-Rabrig[1], Philipp Jugert[2], Birgit Leyendecker[3], Nele McElvany[1]
[1]TU Dortmund, Institut für Schulentwicklungsforschung, [2]Universität Duisburg-Essen, [3]Ruhr-Universität Bochum, Fakultät für Psychologie, Child and Family Research

The transition to secondary school has far reaching consequences for the further educational pathway in the German educational system (e.g., Blossfeld et al., 2019). Achievement development and indicators of the child's achievement (e.g., Dumont et al., 2014) as well as aspects of motivation and well-being (e.g., Steinmayr & Spinath, 2009; Kleinkorres et al., 2020) are central aspects of school success and play an important role for both the teachers' track recommendation and parents' chosen track.

The share of refugee and recent immigrant students (RRI) in Germany has strongly grown in the past years (Sachverständigenrat für Integration und Migration, 2021). Considering that family background can influence various aspects of school success, identifying potentially negative factors and effects is central for this group to counter them. Such differences in family background can influence school success, and transitions especially, via primary background effects transmitted through differences in achievement or other relevant constructs and secondary background effects which are independent of achievement differences (Boudon, 1974).

Learners with immigrant background are underrepresented at academic track schools (Gymnasium; e.g., Autorengruppe Bildungsberichterstattung, 2016). Immigrant background usually relates negatively to school success (e.g., Stanat & Edele, 2015) and negative primary, but not or even positively to secondary effects (e.g., Gresch, 2012). It remains unclear, which background effects can be found for RRI. This group's challenging situation suggest negative family background effects on school success. Furthermore, individual (e.g., vocabulary) and family level factors (e.g. parents' educational aspirations) that can serve as mediators (e.g. Areepattamannil & Lee, 2014; McElvany et al., 2009) are of special interest for this group because of their potentially protective role.

Therefore, the following research questions were regarded: (RQ1) Do patterns of track recommendations differ between RRI, students with second generation, and no immigrant background, controlling for cognitive abilities and socioeconomic status? (RQ2) What role does experience of refuge/immigration, under consideration of other background variables, play for school success (achievement, well-being, track recommendation)? (RQ3) Which individual (educational aspirations child, vocabulary) and family (educational aspirations parents, interest) factors mediate the relations found in RQ2?

The sample included 278 students from 47 classes in the Ruhr area (female: 53.2%; MAge=10.45, SD=0.55), 36.4% were RRI and 39.6% second generation immigrants. The study was conducted in the second half of Grade 4. Standardized tests to measure reading competence, mathematics competence, vocabulary and cognitive abilities, and questionnaires comprising central background variables and psychosocial constructs were applied in the class context or small group. Teachers provided information on track recommendation and grades. Reliabilities were satisfactory ($.61 \leq \alpha \leq .89$).

RQ1 was investigated using logistic regression of track recommendation on immigrant status. For RQ2 we specified a structural equation model with RRI-status, home language, socioeconomic status as independent, and reading competence, math competence, GPA, life satisfaction, track

recommendation as dependent variables. For RQ3, vocabulary, child's and parents' educational aspirations and parental interest were added as mediators.

Compared to students with no/second generation immigrant background, RRI had a lower chance of receiving a Gymnasium track recommendation (OR_noMIG=0.327, p=.018, OR_2.Generation=0.288, p=.001). Considering additional family background variables and controlling for cognitive abilities, RRI-status was related to lower reading competence (β=-.20, p=.008), worse GPA (β=.15, p=.009), and lower track recommendation (β=-.27, p=.002). Individual vocabulary served as a mediator.

Results allow for important insights into RRI's situation at the end of primary school. Negative relations to reading competence, GPA and track recommendation were at least partially mediated by vocabulary, showing the relevance of an early and comprehensive support for vocabulary learning for RRI especially, which can work protective in turn. Implications for research and practice will further be discussed in the presentation.

Refugee and Immigrant Parents' Academic Involvement, Resources, and Children's Adjustment in German Lower-Income Elementary Schools

Sandy Chwastek[1], Julian Busch[1], Nele McElvany[2], Philipp Jugert[3], Birgit Leyendecker[1]
[1]Ruhr-Universität Bochum, Fakultät für Psychologie, Child and Family Research, [2]TU Dortmund, Institut für Schulentwicklungsforschung, [3]Universität Duisburg-Essen

Introduction. Recently arrived refugee and immigrant children's development is influenced by microsystems such as the family and neighborhood (Suárez-Orozco et al., 2018). In Germany, their families likely settle in lower-income, multiethnic neighborhoods, which can put additional risks and barriers on them. Previous studies suggest that refugee (Aghajafari et al., 2019) and immigrant children (Greenfader, 2019; Wong and Schweitzer, 2017) tend to have poorer academic outcomes than non-immigrant children. Their parents show a high willingness to support their children's education at home (Baird, 2015; Gandarilla Ocampo et al., 2020), adapting protective behaviors such as structuring and helping with homework, staying informed, or encouraging their children (Cureton, 2020; Schlaich, 2021). However, recent immigrant parents residing in lower-income neighborhoods are likely to lack resources such as education, language skills, social support, and information on the host country's education system (Gandarilla Ocampo et al., 2020; Rah, Choi & Nguyen, 2009; Schlaich, 2021) and therefore to show a lower school involvement (Cureton, 2020; Zengin & Akdemir, 2020). This study aimed to explore the relations between parental resources, parental academic involvement, and their children's school adjustment among refugee and immigrant parents in low-income, multi-ethnic neighborhoods in Germany.

Methods. We conducted semi-structured interviews with N = 108 parents of third and fourth graders attending schools in lower-income, multiethnic neighborhoods, dividing them into Arabic-speaking refugee families (n = 23), families with two first-generation immigrant parents (n = 28), families with a migration background (at least one parent born in Germany; n = 35), and non-immigrant parents (n = 22). Using scales, we assessed parents' involvement at home and in school, their resources (education, language skills, health, and social support), their knowledge of the education system, educational aspirations, parent-teacher relationship, sense of wellbeing at school (WBS), and children's school adjustment (Strengths and Difficulties Questionnaire, well-being at school, grades). In open-ended questions, we explored the constructs in detail.

Results. Most parents reported several strategies to support their children's education. However, refugee parents involved less in school and had less social support than parents with(out) a migration background. They had the lowest knowledge of the education system and reported higher emotional symptoms in their children than non-immigrant parents. Non-immigrant parents reported poorer health and participated in fewer neighborhood meetings than the other parents. Language skills differed between all groups, with refugee parents having the lowest skills. We also found differences in parents' approaches to education. For example, refugee parents focused more on the quality of lessons when being asked about their WBS than the other groups.

In sum, we found that parents commanding more resources showed more academic involvement and reported fewer socio-emotional problems of their children. Specifically, parents with better language skills and a higher WBS were more involved in school. Parents who placed more emphasis on their children's grades and parents who had more knowledge of the education system read more often with their children. Parents who participated in more neighborhood meetings went on more excursions with their children.

Parents who went on more excursions with their children reported less hyperactivity/inattention in their children. Additionally, parents who had a higher WBS and parents who had a better knowledge of the education system reported fewer emotional symptoms in their children. Parents who had a higher WBS also described their children to have a higher well-being at school. Parents who were more content with their children's grades reported their children to have higher well-being at school and to have better grades.

Conclusion. The results suggest that refugee and immigrant parents' resources are linked to their academic involvement, children's school integration, and well-being in lower-income settings in higher-income countries.

Exploring the school adjustment of refugee and immigrant children: Can the quality of relationships in school make a difference?

Francesca Ialuna[1], Sauro Civitillo[1], Birgit Leyendecker[2], Nele McElvany[3], Philipp Jugert[1]
[1]Universität Duisburg-Essen, [2]Ruhr-Universität Bochum, Fakultät für Psychologie, Child and Family Research, [3]TU Dortmund, Institut für Schulentwicklungsforschung

School adjustment, such as performing well and feeling to belong to school, is an important developmental task that can impact children's future academic trajectories (Motti-Stefanidi et al., 2012). Yet, for refugee and immigrant children in the resettlement phase, school adjustment might be hindered by acculturative challenges, like learning the school language.

From a resilience perspective (Motti-Stefanidi & Masten, 2017), it is important to examine not only acculturative challenges, but also protective factors. Protective factors can mitigate the negative effect of acculturative challenges on children's adjustment (Motti-Stefanidi & Masten, 2017). Such factors are not only present at the individual level (e.g., cognitive abilities), but also in contexts in which children are embedded, like the school. For example, the quality of relationship with teachers and peers can play a protective role and promote children's school adjustment.

The attachment theory (Bowlby, 1969) posits that high quality relationships can enhance children's sense of safety and relatedness. This seems to be particularly relevant for refugee and immigrant

children, as they are in a new and unfamiliar context. According to the attachment and relational perspective (Juang et al., 2018), positive relationships in the new context could therefore promote refugee and immigrant children's adjustment. Accordingly, empirical literature showed that the quality of relationship with teachers (Roorda et al., 2011) and peers (Allen et al., 2018) is positively associated with children's school achievement and belongingness.

Drawing on the resilience (Motti-Stefanidi & Masten, 2017) and the attachment-relational model (Juang et al., 2018), we aim to investigate to what extent school relationship quality protects refugee and immigrant children's school adjustment from the effect of one acculturative challenge, namely school language proficiency.

We hypothesized teacher-child and peer relationship quality to mitigate the negative effect of refugee and immigrant children's low proficiency in the school language (i.e., German) on their school adjustment (i.e., Grade Point Average, reading comprehension, school belongingness).

We collected data from n = 278 elementary school children (Mage = 10.45, SD = .55) in n = 46 classes in Germany. Hereof 36,4% were first-generation, 39.6% were second-generation refugee and immigrant children, and 24% were non-immigrant children. We assessed language proficiency through a German vocabulary test. To measure teacher-child and peer relationship quality, school belongingness and sociodemographic characteristics, we administered a paper-and-pencil questionnaire. Grade Point Average (GPA) was obtained by averaging children's grades in three main subjects. We assessed reading comprehension through a test for elementary school children.

We analyzed data with multilevel regression models. Preliminary analyses showed that teacher-child relationship quality moderated the associations between German proficiency, GPA ($B = .08$, $p < .05$), and reading comprehension ($B = -.37$, $p < .01$). Particularly, we found that higher levels of teacher-child relationship quality mitigated the effect of German proficiency on reading comprehension. Conversely, higher teacher-child relationship quality seemed to heighten the effect of German proficiency on GPA. The moderating effect of teacher-child relationship quality on school belongingness was not significant ($B = -.01$, $p > .05$). As expected, higher levels of peer relationship quality mitigated the effect of German proficiency on school belongingness ($B = -.02$, $p < .05$). Multi-group analyses will be performed to test model differences between children of the first-, second-migration generation and non-immigrant children.

The study expands literature on the importance of the quality of teacher-child and peer relationships by considering its protective role for refugee and immigrant children's school adjustment. School interventions should consider the importance of relationship quality for all children and especially for refugee and immigrant children. Pre-service teacher trainings could focus on strategies how to enhance relationship quality with pupils. Classroom interventions should aim at increasing the relationship quality between classmates.

C-20: Eigenvideografie als Professionalisierungstool für Lehramtsstudierende im Praxissemester

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* R11 T00 D05

Eigenvideografie als Professionalisierungstool für Lehramtsstudierende im Praxissemester

Chair(s): **Verena Zucker** (Westfälische Wilhelms-Universität Münster)

*Diskutant*in(nen):* **Kira Elena Weber** (IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik in Kiel)

Die Videografierung des eigenen Unterrichts gilt als ein wirksames Tool zur Professionalisierung von (angehenden) Lehrpersonen (Gaudin & Chaliès, 2015). Die Wirksamkeit zeigt sich u.a. in Bezug auf die Förderung der professionellen Unterrichtswahrnehmung (Gold et al., 2017; Sherin & van Es, 2009), also der Fähigkeit, lernrelevante Ereignisse im Unterricht erkennen und wissensbasiert analysieren zu können (Junker et al., 2020; Seidel & Stürmer, 2014). Besonderes Potenzial für den Einsatz eigener Videografien scheint aufgrund vielfältiger Analyseanlässe (Drahmann et al., 2018) das universitäre Praxissemester zu bieten. Ein Einsatz in diesem Kontext kann für Studierende jedoch mit Herausforderungen verbunden sein, da die Eigenvideografie als komplexe Aufgabe (Richards et al., 2021) zusätzlich zu vielen weiteren Anforderungen im Praxissemester steht. Systematische Studien, die den Einsatz von eigenen Unterrichtsvideografien ausgehend von den besonderen Rahmenbedingungen des Praxissemesters untersuchen, fehlen weitestgehend. Gleiches gilt für Studien, die die Wirksamkeit des Einsatzes für die Förderung der professionellen Unterrichtswahrnehmung erforschen.

Im Symposium wird ausgehend von den genannten Desideraten die Eigenvideografie als Professionalisierungstool im Kontext des Praxissemesters fokussiert und deren Wirksamkeit sowie Einsatz diskutiert. Leitend sind dabei zwei Fragestellungen:

1) Fördert der Einsatz von Eigenvideografie im Praxissemester die Entwicklung professioneller Unterrichtswahrnehmung?

2) Welche Faktoren beeinflussen die Umsetzung und Analyse der Eigenvideografie im Praxissemester?

Ausgehend von der ersten Fragestellung fokussiert *Beitrag 1* den Einfluss der Eigenvideografie auf die Entwicklung der professionellen Unterrichtswahrnehmung von Klassenführung unter Berücksichtigung der Praxiserfahrung im Rahmen des Praxissemesters. Die Ergebnisse bestätigen dabei grundsätzlich den Effekt von Eigenvideografie auf die Entwicklung der professionellen Unterrichtswahrnehmung von Klassenführung für Praxissemesterstudierende.

Ergänzend wird in *Beitrag 2* die Förderung der professionellen Unterrichtswahrnehmung von Klassenführung bei Praxissemesterstudierenden durch den Einsatz von Eigenvideografie dem Einsatz von Videos unbekannter Lehrpersonen gegenübergestellt. Auch hier bestätigen die Ergebnisse das entsprechende Förderpotenzial der Eigenvideografie. Zusätzlich werden Gemeinsamkeiten und Unterschiede in der Entwicklung zwischen den Gruppen mit eigenen und fremden Videos dargestellt.

Ausgehend von der zweiten Fragestellung wird in *Beitrag 3* der Einfluss persönlicher Voraussetzungen von Praxissemesterstudierenden auf die Bereitschaft sowie Nutzung von Eigenvideografie untersucht. Für beides ergeben sich sowohl förderliche (z.B. Überzeugung zur

Wirksamkeit von Eigenvideografie) als auch hemmende persönliche Voraussetzungen (z.B. zeitliche Belastung) als beeinflussende Faktoren.

In *Beitrag 4* wird dargestellt, welche Emotionen bei Praxissemesterstudierenden in kollaborativen Analysephasen des eigenen Unterrichtsvideos initiiert werden. Dabei zeigen sich drei Typen der Veränderung von Emotionen und drei wesentliche Einflussvariablen auf das emotionale Erleben: Gruppenvertrautheit, Unterstützung durch Personen an der Praktikumsschule im Genehmigungsverfahren zur Videografie und Vorerfahrung mit Videoanalysen.

Beiträge des Symposiums

Welche Rolle spielt die Eigenvideografie bei der Förderung der professionellen Unterrichtswahrnehmung von Klassenführung im Rahmen des Praxissemesters?

Christina Gippert, Robin Junker, Dorothee Seeger, Manfred Holodynski
Westfälische Wilhelms-Universität Münster

Die Fähigkeit, einen professionellen Blick für den eigenen Unterricht zu entwickeln, um lernrelevante Situationen erkennen und diese wissensbasiert interpretieren zu können, wird professionelle Unterrichtswahrnehmung genannt (Stürmer et al., 2013). Sie basiert auf flexibel strukturiertem und elaboriertem Wissen und setzt sich aus den Prozessen des noticing, d.h. der selektiven Wahrnehmung, und des knowledge-based reasoning, d.h. des wissensbasierten Schlussfolgerns bzw. Interpretierens, zusammen (Sherin, 2007). Sie stellt einen Schlüsselaspekt der professionellen Kompetenz von Lehrkräften dar (Blömeke et al., 2015) und sollte dementsprechend schon im Lehramtsstudium gefördert werden.

Verschiedene Studien zeigen einen positiven Effekt der Eigenvideografie auf die Förderung der professionellen Unterrichtswahrnehmung bei praktizierenden Lehrkräften (z. B. Sherin & van Es, 2009). Im universitären Kontext zeigt sich bei Lehramtsstudierenden die Analyse von eigenen Videos der Analyse von Fremdvideos überlegen (Gold et al., 2020, Hellermann et al., 2015), wobei bei den Studien der Einfluss der damit einhergehenden Praxiserfahrung nicht miteinbezogen werden konnte.

Die vorliegende Studie untersucht daher den Einfluss der Eigenvideografie auf die Entwicklung der professionellen Unterrichtswahrnehmung unter Berücksichtigung der Praxiserfahrung im Rahmen des Praxissemesters. Inhaltlicher Schwerpunkt der Wahrnehmungsschulung war die Klassenführung der angehenden Lehrkräfte.

Im Rahmen der aktuell noch laufenden quasi-experimentellen Interventionsstudie mit Prä-Post-FollowUp-Design erhielten Lehramtsstudierende verschiedener Schulformen und -fächer (N = 126) in einer Begleitveranstaltung zum Praxissemester zunächst eine inhaltliche Einführung zu den Themengebieten Klassenführung und professionelle Unterrichtswahrnehmung. Im Anschluss bestand die Möglichkeit, den eigenen Unterricht zu videografieren und diesen anschließend hinsichtlich der Klassenführung zu analysieren (Experimentalgruppe mit Prüfungsleistung, n = 80). Dem gegenüber stand die Kontrollgruppe, die sich nach dem inhaltlichen Einführungsblock gegen eine Prüfungsleistung entschied und im Rahmen des Praxissemesters zwar eigenen Unterricht durchführte, diesen aber weder videographierte noch hinsichtlich Klassenführung analysierte (n = 46). Die Ergebnisse dieser beiden Gruppen werden mit einer unbehandelten Kontrollgruppe (N = 97) verglichen, die ein reguläres universitäres Semester im Lehramtsstudium absolvierte.

Die professionelle Unterrichtswahrnehmung von Klassenführung wurde mittels eines standardisierten geschlossenen Videotests (Gold & Holodynski, 2017), in dem Items zu verschiedenen Unterrichtsclips auf einer vierstufigen Likert-Skala eingeschätzt werden müssen und im Anschluss mit einem Expert*innenrating verglichen werden, erfasst. Zudem wurden mittels Fragebögen das deklarative und konditional-prozedurale Wissen über Klassenführung (Kurz & Lenske, 2022), die klassenführungsbezogenen Selbstwirksamkeitsüberzeugungen (Pfitzner-Eden et al., 2014) sowie klassenführungsbezogene Einstellungen (selbstentwickelte Items) erhoben.

Erste Varianzanalysen mit Messwiederholung zeigen signifikante Unterschiede zwischen den drei untersuchten Gruppen hinsichtlich der Prä-Post-Entwicklung der professionellen Unterrichtswahrnehmung in den drei Klassenführungsfacetten Monitoring ($F(2,220) = 5.378$, $p = .005$, $\eta p2 = .047$), Strukturierung ($F(2,220) = 3.506$, $p = .032$, $\eta p2 = .031$) und Regeln und Routinen ($F(2,220) = 5.661$, $p = .004$, $\eta p2 = .049$). Dabei unterscheidet sich die Entwicklung in der Experimentalgruppe allerdings nur hinsichtlich der Facette des Monitorings signifikant von der Kontrollgruppe ($F(1,124) = 5.908$, $p = .016$, $\eta p2 = .045$).

Beim Wissen über Klassenführung findet sich ein signifikanter Unterschied zwischen den drei untersuchten Gruppen nur hinsichtlich der Entwicklung des deklarativen Wissens ($F(2,177) = 5.350$, $p = .006$, $\eta p2 = .057$), wobei sich Experimental- und Kontrollgruppe nicht signifikant voneinander unterscheiden.

Grundsätzlich bestätigen die Ergebnisse den bereits gefundenen Effekt von Eigenvideografie auf die Entwicklung der professionellen Unterrichtswahrnehmung von Klassenführung, wobei in der vorliegenden Studie durch die Gruppeneinteilung die Praxiserfahrung im Rahmen des Praxissemesters Berücksichtigung fand. Zumindest in Bezug auf das Monitoring scheint nicht die reine Erfahrung von Praxis, sondern die intensive Auseinandersetzung mit dieser entscheidend zu sein.

Die nicht erwartete positive Entwicklung der Kontrollgruppe in den Facetten Strukturierung und Regeln und Routinen der professionellen Unterrichtswahrnehmung sowie im Wissen über Klassenführung ist sowohl durch den inhaltlichen Input als auch durch die Planung und Durchführung eigenen Unterrichts im Praxissemester erklärbar.

Zur Förderung professioneller Unterrichtswahrnehmung im Praxissemester anhand von Eigen- und Fremdvideos

Verena Oestermann, Ulrike Weyland, Wilhelm Koschel
Westfälische Wilhelms-Universität Münster

Theoretischer Hintergrund

Die Fähigkeit zur professionellen Unterrichtswahrnehmung gilt als wesentlicher Bestandteil von Lehrer*innenprofessionalität (Blömeke et al., 2015; Kunter et al., 2013). Um diese Fähigkeit bei angehenden Lehrpersonen bereits während des Studiums zu fördern, wird zunehmend auf das Potenzial von Unterrichtsvideos, auch im Zusammenhang mit Praxisphasen, gesetzt. Die Wirkung von Unterrichtsvideos zur Förderung professioneller Unterrichtswahrnehmung wird seit einiger Zeit in interdisziplinären Projektgruppen an der WWU Münster in unterschiedlichen Kontexten erforscht

(Koschel & Weyland, 2020a; Gold et al., 2017). Je nach hochschuldidaktischer Einbindung der Videos kann von einem positiven Lerneffekt ausgegangen werden (Koschel & Weyland, 2020b).

Derartige Befunde beziehen sich jedoch bislang vorrangig auf den Einsatz von fremden Unterrichtsvideos. Die Forschung zu Effekten der Analyse eigener Unterrichtsvideos, denen ebenfalls ein hohes Potential zur Förderung professioneller Unterrichtswahrnehmung zugeschrieben wird (Gold et al., 2017; Hellermann et al., 2015; Seidel et al., 2011), ist bislang aber deutlich unterbelichtet. Langzeitpraktika, in denen die Einnahme und Ausbildung einer for-schenden Grundhaltung als Zielperspektive intendiert wird, bieten zur Erforschung eben genannter Effekte einen geeigneten Rahmen (Weyland, 2014).

Fragestellung

In der vorliegenden Studie wird daher der Frage nachgegangen, inwiefern sich die professionelle Unterrichtswahrnehmung Studierender im Praxissemester hinsichtlich klassenführungsrelevanter Ereignisse durch den Einsatz von Eigen- und Fremdvideos unterscheidet.

Methode

In einem quasi-experimentellen Prä-Post-Kontrollgruppen-Design (EG1: Analyse von Eigenvideos, EG2: Analyse von Fremdvideos, KG: keine Videoanalyse) wird die professionelle Unterrichtswahrnehmung Studierender verschiedener Lehrämter (erwartetes N = 129) erfasst. Die Intervention findet im Rahmen der Begleitveranstaltungen zum Praxissemester statt und erstreckt sich über drei Erhebungszeitpunkte. Nach einer umfassenden An- und Begleitung analysieren die Studierenden der EG1 und EG2 regelgeleitet ihr eigenes bzw. ein fremdes authentisches Unterrichtsvideo aus der Perspektive einer effektiven Klassenführung.

Innerhalb der Prä-/Posttests kommt ein offenes Videoanalyseinstrument zum Einsatz, das sich an den Ansätzen LAF und VAST orientiert (siehe hierzu Koschel, 2021). Das verwendete Testvideo stammt aus einem beruflichen Unterricht und wurden durch ein Masterrating validiert. Die studentischen Analysen werden inhaltsanalytisch kodiert und varianzanalytisch ausgewertet.

Ergebnisse und ihre Bedeutung

Die Ergebnisse des Prä-Post-Tests des ersten Erhebungszeitraums zeigen, dass sich die Studierenden von EG1 und EG2 im Vergleich zur KG signifikant in ihrer professionellen Unterrichtswahrnehmung verbessern. Dabei ist auffällig, dass die EG1 ein höheres Einstiegsniveau für den Analyseschritt „Interpretation“ aufweist als die EG2. Ein möglicher Erklärungsansatz liegt in der angenommenen größeren Bereitschaft der EG1 an der eigenen Professionalisierung zu arbeiten. Für den Analyseschritt „Bewertung“ ist hingegen ein Abfall der EG1 im Vergleich zur EG2 zu konstatieren. Möglicherweise wird den Studierenden der EG1 während der Analyse ihrer eigenen Unterrichtsvideos bewusst, wie schwierig die Bewertung des eigenen Handelns ist, insbesondere dann, wenn Handlungsabläufe nicht gelungen sind, sodass sie dann in der Bewertung der Lehrkraft des Testvideos nachsichtiger sind. Im Rahmen des Vortrags werden der aktuelle Auswertungsstand präsentiert und mögliche Implikationen für die Lehrer*innenbildung diskutiert.

Was motiviert oder hemmt Studierende, sich im Praxissemester videografieren zu lassen? Eine Untersuchung zur Bedeutung persönlicher Voraussetzungen für die Nutzung von Eigenvideografie

Verena Zucker, Nicola Meschede
Westfälische Wilhelms-Universität Münster

Die Analyse eigener Unterrichtsvideos gilt als eine wirksame Möglichkeit, um Kompetenzen wie z.B. die professionelle Unterrichtswahrnehmung oder Selbstwirksamkeitsüberzeugungen zu fördern (Gold et al., 2017). Während Studien zeigen, dass videobasierte Selbstreflexionen mit diversen Hürden für (angehende) Lehrkräfte verbunden sein können (Kleinknecht & Schneider, 2013; Pollmeier et al., 2022), wurden förderliche und hemmende Faktoren für die Nutzung von Eigenvideografie bislang kaum systematisch untersucht (Leung et al., 2019). In Anlehnung an Befunde der Fort- und Weiterbildungsforschung (E. Richter et al., 2018; Scanlan & Darkenwald, 1984) können hierbei u.a. persönliche Voraussetzungen als bedeutsam angesehen werden. Während z.B. die Überzeugung zur Wirksamkeit des Angebots (Warketin, 2018) oder die eigene Entwicklungsorientierung (Rzejak et al., 2014) als förderliche Faktoren angenommen werden, wurden als hemmende Faktoren das Disengagement (E. Richter et al., 2018) oder negative Emotionen (Pollmeier et al., 2022) identifiziert.

Die vorliegende Studie hat davon ausgehend das Ziel, den Einfluss persönlicher Voraussetzungen auf die (Nicht-)Nutzung von Eigenvideografie im Praxissemester zu untersuchen. Dieses Ziel konkretisiert sich in zwei Fragestellungen:

(1) Welche persönlichen Voraussetzungen von Studierenden hemmen bzw. fördern die grundlegende Bereitschaft zur Eigenvideografie im Praxissemester?

(2) Welche persönlichen Voraussetzungen von Studierenden hemmen bzw. fördern die tatsächliche (Nicht-)Nutzung von Eigenvideografie im Praxissemester?

Im Rahmen der Studie wurden insgesamt 217 Masterstudierende (Semester: M = 8.85, SD = 2.53; Alter: M = 24.45, SD = 2.45) unterschiedlicher Lehrämter vor Beginn des Praxissemesters online befragt. Dabei wurden die grundsätzliche Bereitschaft zur Eigenvideografie sowie potenziell förderliche (Überzeugung zur Wirksamkeit von Eigenvideografie, Selbstwirksamkeitsüberzeugungen bzgl. der eigenen Videoanalyse, Interesse an der eigenen Weiterentwicklung, Reflexionsbereitschaft, technische Fähigkeiten, Wissen zum Datenschutz, Erfahrung mit Eigenvideografie) sowie hemmende persönliche Voraussetzungen (Misserfolgsbefürchtungen, zeitliche Belastung, Disengagement) u.a. in Anlehnung an bestehende Instrumente (Lenski et al., 2016; Kunter et al., 2019) erfasst. Die Einschätzung von Items erfolgte auf einer vierstufigen Likert-Skala (1 = trifft nicht zu – 4 = trifft voll zu) bzw. bei der Erfahrung mit Eigenvideografie dichotom (0 = keine Erfahrung, 1 = Erfahrung). Die (Nicht-)Nutzung der Eigenvideografie wurde nach dem Praxissemester erfragt.

(1) Eine multiple lineare Regression zeigt, dass die Überzeugung zur Wirksamkeit von Eigenvideografie ($b = 0.92$, $SE = 0.12$, $p < .001$), die technischen Fähigkeiten ($b = -0.11$, $SE = 0.46$, $p = .016$) sowie die Misserfolgsbefürchtungen ($b = -0.34$, $SE = 0.06$, $p < .001$) einen Einfluss auf die Bereitschaft zur eigenen Videografierung vor dem Praxissemester haben, ($F(10, 206) = 18{,}43$, $p < .001$). Insgesamt wird durch diese Faktoren 45 % der Streuung in der Bereitschaft erklärt.

(2) Eine multiple logistische Regression zeigt zudem, dass die Nutzung der Eigenvideografie von der Überzeugung zur Wirksamkeit von Eigenvideografie (b = 0.96; OR = 2.61, SE = 0.50; p = .05), den Selbstwirksamkeitsüberzeugungen hinsichtlich der eigenen Videoanalyse (b = 0.69; OR = 1.99, SE = 0.33; p = .04), den technischen Fähigkeiten (b = -0.41; OR = 0.66, SE = 0.20; p = .034), der zeitlichen Belastung (b = -0.51; OR = 0.60, SE = 0.24; p = .03) und der Erfahrung mit Eigenvideografie (b = -1.35; OR = 0.26, SE = 0.49; p = .006) beeinflusst wird, $\chi^2(11) = 42{,}192$, $p < .001$. Die Varianzaufklärung ist akzeptabel (Nagelkerkes $R^2 = .236$).

Die Ergebnisse geben Aufschluss über verschiedene förderliche sowie hemmende Faktoren für die Bereitschaft sowie Nutzung von Eigenvideografie (E. Richter et al., 2018; Scanlan & Darkenwald, 1984). Die geringere Aufklärung der (Nicht-)Nutzung im Vergleich zur Bereitschaft könnte darauf zurückzuführen sein, dass neben persönlichen Voraussetzungen auch weitere Faktoren wie der Kontext (z.B. schulische Bedingungen) oder die Qualität der Lerngelegenheit (z.B. das Praxissemesterbegleitseminar) die (Nicht-)Nutzung beeinflusst haben könnten (D. Richter et al., 2019). Um das Potential der Eigenvideografie in der Lehrkräftebildung langfristig ausschöpfen zu können, sollten die identifizierten Faktoren gezielt berücksichtigt werden.

Zwischen Angst und Vorfreude – Emotionales Erleben Lehramtsstudierender bei der Arbeit mit Eigenvideografien

Pascal Pollmeier[1], Christoph Vogelsang[1], Tim Rogge[2]
[1]Universität Paderborn, [2]Qualitäts- und Unterstützungsagentur - Landesinstitut für Schule (QUA-LiS) NRW

Unterrichtsvideografien werden in der Lehramtsausbildung zunehmend als methodisches Mittel zur Förderung reflexiver Kompetenzen implementiert. Dabei konnten überwiegend positive Effekte nachgewiesen werden (Gaudin & Chaliès, 2015). Es wird aber angenommen, dass die Arbeit mit Eigenvideos im Vergleich zur Arbeit mit Fremdvideos mit negativeren Emotionen einhergeht (Kleinknecht & Poschinski, 2014), insbesondere bei kollaborativer Reflexion, die für die Implementationen im Lehramtsstudium typisch ist (z.B. Herbst, 2019). Emotionen sind eine komplexe Reaktion auf die kognitive Einschätzung einer spezifischen Situation (Moors, Ellsworth, Scherer & Frijda, 2013). Sie werden typischerweise als negativ oder positiv (Valenz) erlebt. Im Kontext der Hochschullehre entstehen negative Emotionen, wenn eine Lernaktivität und ihre erwarteten Folgen negativ bewertet werden und zugleich der Eindruck geringer Kontrolle über ihren Ausgang besteht (Pekrun, 2006). Negative Emotionen haben dabei einen negativen Einfluss auf Lernprozesse (z.B. Kordts-Freudinger et al, 2017). Vor diesem Hintergrund kann erwartet werden, dass bei der kollaborativen Reflexion von Eigenvideografien vermehrt negative Emotionen entstehen (Yerrick et al., 2005), bspw. durch die Erwartung eines Gesichtsverlusts, was wiederum in Vermeidungsverhalten (Leung et al., 2021) oder geringer Akzeptanz der Eigenvideografie als Methode resultieren kann. Emotionen bzgl. der Arbeit mit Eigenvideografien im Lehramtsstudium wurden bisher allerdings wenig untersucht.

Es wurde daher explorativ analysiert welche Emotionen bei Lehramtsstudierenden initiiert werden, die im Praxissemester Eigenvideografien von Unterricht kollaborativ mit Kommiliton*innen reflektieren (müssen) und wie sich diese im Verlauf verändern. Hierzu wurden N=20 Lehramtsstudierende (15 LA HRSGe, 5 LA GyGe, 2. Mastersemester) in einer längsschnittlichen Interviewstudie herangezogen,

die im Praxissemester in einer universitäre Begleitveranstaltung Eigenvideografien erstellt und gemeinsam mit Kommiliton*innen reflektiert haben. Um die Emotionen der Studierenden und ihre Veränderungen zu erfassen, wurden strukturierte Prä-Post-Interviews geführt. Die Prä-Interviews fokussierten auf Vorerfahrungen, den emotionalen Zustand bzgl. der kommenden Eigenvideografie sowie der Antizipation der Reflexionssituation. Im Post-Interview wurden die emotionalen Zustände bei der Vorbereitung, Durchführung und Reflexion der Videografie betrachtet und ins Verhältnis zu anderen Feedbacksituationen gesetzt. Die Auswertung erfolgte qualitativ-inhaltsanalytisch (Kuckartz, 2016), wobei nicht nur explizite emotionale Äußerungen, sondern auch indirekte betrachtet wurden (Hailikari et al., 2016). Anschließend folgte eine „reflektierende Interpretation" (Bohnsack 2018), in welcher das Entstehen emotionaler Zustände vor dem Hintergrund von Rahmenbedingungen und sozialer Prozessstrukturen betrachtet wurde.

Insgesamt konnten drei Typen der Veränderung von Emotionen bei der kollaborativen Arbeit mit Eigenvideografien rekonstruiert werden: Typ 1 mit Veränderung von zu Beginn negativen emotionalen Zuständen zu positiven Zuständen nach der Videografieerfahrung. Typ 2 mit stabil positiven Emotionen und Typ 3 umgekehrt mit stabil negativen Emotionen. Ein Muster von positiven zu negativen Emotionen wurde nicht beobachtet. Übergreifend ergaben sich drei wesentliche Einflussvariablen auf das emotionale Erleben. Gruppenvertrautheit reduziert das Erwarten negativer Folgen im Sinne kollektiver Kontrolle, da innerhalb der Reflexionsphase weniger mit negativen Zuschreibungen zur eigenen Person zu rechnen ist. Ein weiterer Einflussfaktor ist die Unterstützung durch Personen an der Praktikumsschule im Genehmigungsverfahren zur Videografie. Fehlt diese bzw. wird Videografie erschwert, entsteht ein Konflikt der Studierenden bei der Bewältigung widerstrebender Anforderungen von Universität und Schule. Vorerfahrung mit Videoanalysen konnte als starker Prädiktor positiven emotionalen Erlebens beobachtet werden, insbesondere im Vergleich von Typ 2 und 3, wobei ersterer positive, zweiterer negative Vorerfahrungen (wohlgemerkt in anderen Kontexten) berichtet. Die Analysen zeigen insgesamt, dass weniger die Tatsache, dass Mitstudierende das eigene Video betrachten, der Auslöser negativer Emotionen ist, sondern eher ein individueller negativer „evaluative stance " (vgl. Leung et al., 2021). Die Studierenden äußern fallübergreifend unterschiedliche Emotionen, je nachdem welche Statusgruppe Feedback zur Eigenvideografie gibt. Fachleiter*innen werden als hierarchisch höherstehend und mit der Erwartung eines primär bewertenden Feedbacks assoziiert (entgegen den Zielsetzungen des Praxissemesters), während das Feedback anderer Gruppen (z.B. schulische Mentor*innen) primär als beratend, weniger emotional belastend, aber auch weniger wertvoll erwartet wird. Aus den Befunden werden Vorschläge für die Gestaltung von Eigenvideografiearbeit abgeleitet.

Sonderveranstaltung

C-21: Sonderveranstaltung: DFG

Zeit: Mittwoch, 01.03.2023: 10:30 - 12:15 · *Ort:* R14 R02 B07

DFG-Förderung in der Erziehungswissenschaft und Bildungsforschung: Aktuelle Informationen zu Förderprogrammen, Antragstellung und Begutachtungs- und Entscheidungsverfahren

Annabell Zentarra
Deutsche Forschungsgemeinschaft, Deutschland

Die Deutsche Forschungsgemeinschaft (DFG) ist die zentrale Selbstverwaltungseinrichtung der Wissenschaft zur Förderung der Forschung an Hochschulen und öffentlich finanzierten Forschungsinstitutionen in Deutschland. Im Jahr 2022 hat die DFG Forschungsvorhaben im Bereich Erziehungswissenschaft und Bildungsforschung mit ca. 11 Mio. Euro gefördert und ist damit für Forschende eine wesentliche Quelle für Drittmittel.

Im Rahmen der Veranstaltung werden aktuelle Informationen zu den Förderprogrammen der DFG, zur Antragstellung und zu den Begutachtungs- und Entscheidungsverfahren präsentiert. Welche Forschungsvorhaben können mit welchen Förderinstrumenten unterstützt werden? Welche Formalia sollten bei der Antragstellung beachtet werden? Wann brauche ich ein Ethikvotum für die Antragstellung? Wie sollte ein „guter“ Antrag aufgebaut sein? Wie können nationale und internationale Kooperationen umgesetzt werden? Nach welchen Kriterien wird begutachtet? Wer sind die Gutachterinnen und Gutachter? Wie sieht der weitere Entscheidungsprozess aus? Zudem besteht die Möglichkeit, mit der Vertreterin der DFG-Geschäftsstelle über Fragen und Probleme bezüglich der Einwerbung von DFG-Mitteln zu diskutieren.

Session D

Einzelbeiträge

D-01: COVID-19 Pandemie II

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S06 S00 B29
Chair der Sitzung: Hedwig Gasteiger

COVID-19-bedingte Schulschließungen und mathematische Leistungen - Ergebnisse einer wiederholten Querschnittsstudie in Jahrgangsstufe 3

Hedwig Gasteiger[1], Karoline A. Sachse[2], Kristoph Schumann[2], Mona Gerve[1], Axel Schulz[3], Maria Engelbert-Kocher[2]
[1]Universität Osnabrück, Germany; [2]Institut für Qualitätsentwicklung im Bildungswesen an der Humboldt-Universität zu Berlin, Germany; [3]Universität Bielefeld, Germany

Theoretischer Hintergrund

Aufgrund der COVID-19-Pandemie wurden ab dem Frühjahr 2020 Maßnahmen ergriffen, die in Deutschland bundeslandspezifisch bis zum Sommer 2021 zu Schulschließungen, Wechselunterricht, kurzzeitigen Schulöffnungen und erneuten Phasen des Distanzunterrichts führten (Jungblut, 2020; Niedersächsisches Kultusministerium, 2022). Zu dieser herausfordernden Schulsituation gibt es v.a. Ergebnisse aus Befragungen von Lehrkräften, Eltern und Lernenden: Ein knappes Drittel unterrichtete digital (Helm et al., 2021). Die individuelle Lernzeit wurde mit nur 2h bis 5h täglich angegeben (Helm et al., 2021). Die Zeit im Distanzlernen wurde v.a. für die Bearbeitung von Arbeitsblättern verwendet und es gab kaum Möglichkeit für Austausch im Klassenverband (Wößmann et al., 2020). Offensichtlich fand weniger kognitive Aktivierung statt (Helm et al., 2021), es wurde eher transmissiv unterrichtet (Aldon et al., 2021) und der Schwerpunkt lag eher auf prozeduralem Wissen (Rechnen, Fertigkeiten). Bislang gibt es einige Ergebnisse aus Vergleichsstudien zu den Auswirkungen der veränderten Schulsituation auf mathematische Leistungen. Schult et al. (2022) zeigten 2020 einen Lernrückstand (1/5 Schuljahr) im Vergleich zu den 3 Jahren zuvor (Jgst. 5, Baden-Württemberg). Ein Leistungsabfall wurde auch in der Primarstufe festgestellt (Deutschland: Stanat et al., 2022; Schweiz: Tomasik et al., 2021; Niederlande: Engzell et al., 2021). Kaum Leistungsunterschiede berichten Depping et al. (2021) aus ihrer Hamburger Vergleichsstudie (2019 vs. 2020; Leitideen *Daten, Häufigkeit, Wahrscheinlichkeit*; *Raum und Form)*.

Fragestellung

Laut den Befragungsergebnissen veränderte sich Unterricht 2020/2021 gravierend. Vor allem für den deutschen Kontext gibt es wenige vergleichende Studien, die die mathematische Leistung in der Primarstufe detailliert untersuchen. Deshalb widmeten wir uns folgenden Fragen:

- Zeigt sich ein Leistungsrückgang in Mathematik zwischen 2019 und 2021 bei Kindern in Jahrgangsstufe 3?
- Zeigt sich ein unterschiedlicher Leistungsrückgang in verschiedenen Leitideen bzw. Anforderungsbereichen?

Methode

Im Rahmen der Pilotierungsstudien der bundesweiten Vergleichsarbeiten VERA-3 wurden 2019 Daten von *N*=1905 und 2021 von *N*=3203 erhoben. Die Kinder bearbeiteten 551 Items zu allen fünf Leitideen der Bildungsstandards (71 Items identisch in beiden Jahren). Mit Hilfe eines linearen logistischen Testmodells mit zusätzlichem Fehlerterm im Framework der Generalized Linear Mixed Models mit einer Mehrebenenstruktur wurden die Auswirkungen des Testzeitpunkts, der inhaltlichen

Leitideen und der Anforderungsbereiche auf die korrekten Antworten der Kinder untersucht (Haupt- und Interaktionseffekte).

Ergebnisse und Diskussion

Es zeigte sich ein negativer Haupteffekt des Testzeitpunkts auf alle Items (p<.001), das heißt, dass Items im Mittel (bezogen auf Anforderungsbereiche, Leitideen, Kompetenzstufen) 2021 eine deutlich geringere Lösungshäufigkeit aufwiesen als 2019. Mit Ausnahme von *Muster und Strukturen* wurden Items aller Leitideen 2021 signifikant seltener gelöst als 2019. Der Leistungsverlauf bei Items zu *Zahlen und Operationen* war ähnlich zum durchschnittlichen Verlauf aller Items. Bei *Raum und Form*und *Daten, Häufigkeit, Wahrscheinlichkeit* zeigte sich ein deutlich stärkerer Abfall.

Zu beiden Testzeitpunkten gab es wenige Items aus Anforderungsbereich III. Deshalb wurde verglichen, ob sich die Bearbeitung von Items mit Anforderungsbereich I (Reproduzieren) unterschied von Items mit Anforderungsbereich II/III (Zusammenhänge herstellen, Verallgemeinern). Es zeigte sich ein Interaktionseffekt mit dem Erhebungsjahr (p<.05). Aufgaben zu Anforderungsbereich II/III zeigten einen signifikant stärkeren Abfall.

Gründe für die signifikanten Unterschiede zwischen 2019 und 2021 könnten in der berichteten verringerten Lernzeit liegen und/oder im offensichtlich weniger kognitiv aktivierend gestalteten Unterricht. Die Leitideen *Raum und Form* und *Daten, Häufigkeit, Wahrscheinlichkeit* wurden 2021, anders als bei Depping et al. (2021), deutlich schlechter bearbeitet als 2019. Der Unterschied in den Ergebnissen könnte daran liegen, dass Corona-Aufholprogramme, die in Hamburg offensichtlich gut genutzt wurden, zwischen den Erhebungen bei Depping et al. (2021) lagen. Zudem ist denkbar, dass durch die eher prozedural ausgerichtete Art des Unterrichtens und ein intensives Bearbeiten von Arbeitsblättern die für den Verständnisaufbau wichtigen geometrischen Handlungsaktivitäten und Zufallsexperimente in der COVID-19 bedingt veränderten Unterrichtssituation kaum realisiert wurden. Die prozedural ausgerichtete Arbeit erklärt auch den signifikant unterschiedlichen Verlauf der Lösungsquoten in den beiden Anforderungsbereichsgruppen.

Two pandemic years greatly reduced young people's life satisfaction: Evidence from Comparison of Pre- and Post-COVID-19 Panel Data.

Martin Neugebauer[1], Alexander Patzina[2], Hans Dietrich[3], Malte Sandner[4]
[1]Freie Universität Berlin; [2]Otto-Friedrich-Universität Bamberg; [3]Institut für Arbeitsmarkt- und Berufsforschung (IAB); [4]Technische Hochschule Nürnberg

The COVID-19 pandemic and the restrictions to prevent its spreading have turned adolescents' lives upside down. Young people aged 16-25 went into social isolation (e.g., Elmer et al., 2020), and they were faced not only with learning losses (Betthäuser et al., 2022) and an erosion of their daily structure (e.g., Grewenig et al., 2021), but also with insecurity concerning their educational and occupational future (e.g., Aucejo et al., 2020). While everyone was (and still is) affected by the pandemic, young people are in a particularly sensitive developmental stage, which has led to concerns about their well-being (Kowal et al., 2020; OECD, 2022).

A rapidly emerging literature examines the well-being of adolescents during the pandemic (e.g., Jensen & Reimer, 2021; Magson et al., 2021; Ravens-Sieberer et al., 2022; Sandner et al., 2022; Thorisdottir et al., 2021; review: Zolopa et al., 2022), indicating an overall negative impact of the

pandemic. However, these important contributions are typically faced with several challenges: First, they draw on survey data collected during the pandemic, and often lack a control group, which was unaffected by the pandemic. Without a control group, the estimation of causal effects is notoriously difficult. Second, the bulk of research describes merely immediate changes in well-being during the initial lockdown in spring 2020, and it remains unclear if the pandemic triggered longer lasting consequences for the well-being of adolescents. Third, while several studies point to more detrimental effects for girls (Zolopa et al., 2022), much less is known about heterogeneous effects of the pandemic with respect to students' personality traits.

The aim of this paper is to contribute to the literature by addressing these issues. To identify causal effects, we compare trends in life satisfaction of high school students affected by the pandemic, with two pre-pandemic cohorts of high school students in a difference-in-differences design. Data from the 'treatment' cohort comes from a subsample of the IAB BerO panel study, in which N = 2374 high school students were repeatedly surveyed during the last two years of high school and 6 months after finishing their 'Abitur' (in 2019, 2020, 2021). Data from the control cohorts are taken from the NEPS starting cohorts 3 (N = 1433) and 4 (N = 3420), who graduated high school in 2014 and 2018, respectively. The data allow us to measure trends in life satisfaction two years into the pandemic. They also allow us to measure whether young people responded differently to the pandemic based on their personality profile.

First results show a substantial decline in life satisfaction over the course of the pandemic, much stronger than comparative estimates from the general German population. Life satisfaction remained much lower even two years into the pandemic, compared to pre-pandemic control groups. Extraverted and emotionally stable people showed the strongest decline in life satisfaction.

Veränderungen in der Schulqualität während COVID-19. Ergebnisse einer Längs- und einer Querschnittsstudie mit Lehrkräften.

Christoph Helm[2], **Stephan Gerhard Huber**[1]
[1]Pädagogische Hochschule Zug, Schweiz; [2]Johannes Kepler Universität Linz, Österreich

Die COVID-19-Pandemie führte weltweit zu wiederholten Schulschließungen. Viele Studien belegen die negativen Auswirkungen von Schulschließungen auf die Lernzeit der Schüler:innen, ihre Leistungen und ihr Wohlbefinden. Dagegen ist wenig darüber bekannt, wie sich die Krise auf Aspekte der Schulqualität (z. B. die Unterrichtsqualität) auswirkt. Einbußen in der Schulqualität sind jedoch erwartbar, da die Pandemie die Schulen völlig unvorbereitet traf. Dementsprechend hatten die Lehrkräfte keine Zeit, sich auf die Umstellung von regulärem Unterricht auf Fernunterricht und Online-Lernen vorzubereiten. Die Forschung (vgl. Bremm et al., 2021) zeigt, dass die wichtigste Ressource zur Reaktion auf die Herausforderungen des Fernunterrichts die Zusammenarbeit der Lehrkräfte war (z. B. die gegenseitige Unterstützung). Sie stelle eine zentrale Schulentwicklungskapazität in Krisenzeiten dar. Die Verlagerung vom Präsenzunterricht zum Fernunterricht machte die Koordination und Orchestrierung von Lernangeboten innerhalb der einzelnen Schulen wesentlich (vgl. Bremm et al., 2021).

Vor diesem Hintergrund zielt unsere Studie darauf ab, Veränderungen in der wahrgenommenen Schulqualität während der Pandemie zu untersuchen und zu erforschen, ob die Zusammenarbeit der Lehrer:innen Veränderungen in der Schulqualität vorhersagt.

H1: Im Verlauf der Pandemie gehen wir von Verlusten bei Dimensionen der Schulqualität aus, für die der Fernunterricht eine Herausforderung darstellt (bspw. Unterrichtsqualität, Qualität der Schulleitung).

H2: Eine höhere Qualität der Zusammenarbeit von Lehrer:innen steht in Zusammenhang mit geringeren Verlusten in den Dimension der Schulqualität.

H3: Höhere Ausgangswerte in der Schulqualität sind mit geringeren Verlusten verbunden ("Matthäus-Effekt").

Methode

Um unsere Hypothesen zu testen, verwenden wir Daten aus zwei separaten Studien.

Studie 1 - Längsschnittdaten

Daten zur Schulqualität wurden wiederholt (3 Zeitpunkte) von 2.627 Lehrkräften an 120 Schulen in Nordrhein-Westfalen erhoben. Während 88 Schulen in den Jahren 2016/17-2018/19 untersucht wurden (= Kontrollgruppe), liegen für 32 Schulen Daten für die Zeit vor und nach den COVID-19-bedingten Schulschließungen, d.h. in den Jahren 2018-2020, vor (= Corona-Gruppe).

Studie 2 - Querschnittliche Daten

Es wurde eine Online-Befragung, vom 1. Dezember 2020 bis zum 9. August 2021 bei 4.113 Lehrkräften durchgeführt wurde. Die Teilnehmenden wurden gebeten, Merkmale der Schulqualität für unterschiedliche Phasen der Pandemie rückblickend zu bewerten. Darüber hinaus wurden die Teilnehmenden gebeten, die gegenseitige Unterstützung der Lehrkräfte an ihrer Schule vor der COVID-19-Pandemie zu schätzen.

Mit Hilfe von Latent-Growth-Curve-Modellen testen wir in beiden Studien die oben genannten Hypothesen.

Ergebnisse

H1. Überraschenderweise zeigen die Ergebnisse der Längsschnittstudie (Studie 1), dass die Lehrkräfte der Corona-Gruppe eine bessere Entwicklung der Schulqualitätsaspekte berichten als die Lehrkräfte der Kontrollgruppe. Im Gegensatz dazu bestätigen die Ergebnisse der Querschnittsstudie (Studie 2) unsere Annahme, dass Aspekte der Schulqualität durch die Pandemie negativ beeinflusst wurden.

H2. Überraschenderweise korreliert die Lehrerkooperation mit höheren Einbußen bei der Schulqualität während der Pandemie.

H3. In beiden Studien sind die Korrelationen zwischen den Ausgangswerten und den Entwicklungen der Schulqualität negativ, d. h. in Schulen mit höherer Qualität vor der Pandemie waren die Verluste bei den Schulqualitätsaspekten überraschenderweise höher und die Gewinne bei den Schulqualitätsaspekten niedriger.

Diskussion

Die überraschenden Ergebnisse werden vor dem Hintergrund möglicher Erklärungen (theoretische und methodische Gründe), der Einschränkungen der Studie (bspw. Einzelebenenanalyse, Lehrkräfteeinschätzungen, retrospektive Bewertung von Veränderungen) und der praktischen Implikationen, z. B. der Relevanz der Lehrkräftekooperation als zentrale Schulentwicklungskapazität in Krisenzeiten, diskutiert.

Emotionale Erschöpfung von Lehrkräften während der Covid-19 Pandemie: Analyse individueller und schulischer Prädiktoren

Tabea Schulze-Hagenest, Uta Klusmann, Bastian Carstensen, Thorben Jansen, Jennifer Meyer, Olaf Köller
IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany

Emotionale Erschöpfung stellt das Kernsymptom Burnouts dar (Maslach et al., 2001; Schaufeli & Taris, 2014) und ist aufgrund der vielfältigen Verantwortungen im Lehrerberuf von besonderem Forschungsinteresse. So hat Burnout bei Lehrkräften vielfältige negative Konsequenzen sowohl für die Gesundheit der Lehrkräfte selbst als auch für die Schülerinnen und Schüler sowie die Schulen (Arens & Morin, 2016; Klusmann et al., 2021; Wartenberg et al., submitted). Bereits vor der Covid-19 Pandemie gaben Lehrkräfte im Vergleich zu anderen Berufsgruppen eine höhere Belastung an (Iriarte Redín & Erro-Garcés, 2020; Johnson et al., 2005) und aktuelle Befunde weisen darauf hin, dass Lehrkräfte infolge der Pandemie nochmals stärkere Burnout-Symptome berichten (Alves et al., 2020; Ozamiz-Etxebarria et al., 2021).

Zwei zentrale Theorien, die zur Erklärung der Entwicklung von Burnout herangezogen werden können, betonen einerseits die Bedeutung des Kontextes sowie der spezifischen Charakteristiken des Arbeitsumfeldes (Job Demands-Resources Model; Bakker & Demerouti, 2007) und andererseits die individuelle Bewertung von Ressourcen und Stressoren als zentralen Faktor (Transaktionale Stressmodell; Lazarus & Folkmann, 1984). Dabei konnten in bisherigen Studien bereits vielfältige soziale (z. B. Unterstützung durch das Kollegium) und individuelle (z. B. Persönlichkeit) Faktoren zur Erklärung von Unterschieden im Belastungserleben der Lehrkräfte identifiziert werden (z. B. Collie, 2021; Kim et al., 2021; Klusmann et al., submitted). Jedoch ist bislang wenig darüber bekannt welche Rolle Merkmale auf Schulebene für das Erleben von Burnout-Symptomen während der Covid-19 Pandemie spielen (Klusmann et al., 2008, Van Droogenbroeck et al., 2021). Insgesamt berücksichtigen bisherige Befunde vor allem die individuelle Bewertung des Arbeitsumfeldes, die sich zwischen den Personen stark unterscheiden kann. Um den Kontext adäquat zu erfassen, können zum einen objektive Kriterien oder externe Angaben herangezogen werden und zum anderen können individuelle Angaben in Form aggregierter Werte berücksichtigt werden.

Die aktuelle Studie untersucht daher die folgenden zwei Forschungsfragen:

1. Wie stark sind die Unterschiede in der emotionalen Erschöpfung der Lehrkräfte sowie in den Schulcharakteristiken zwischen Schulen während der Covid-19 Pandemie?
2. Inwiefern können diese Differenzen und die Unterschiede zwischen Personen durch Schulcharakteristiken und individuelle Faktoren – Ressourcen (z. B. Unterstützung durch die Schulleitung) und Stressoren (z. B. sozioökonomischer Hintergrund der Schülerinnen und Schüler) – erklärt werden?

Die Datenbasis bilden die Angaben von 1024 Lehrkräften aus 120 Schulen, die im Frühjahr 2021 online befragt wurden. Zudem finden Antworten von 155 Schulleitungen als Einschätzungen auf Schulebene Berücksichtigung. Die Fragestellungen wurden mittels Mehrebenenanalysen untersucht.

Dabei wurde ein Random Intercept Modell gewählt, bei dem die Variablen auf dem individuellen Level Group-mean zentriert aufgenommen wurden.

Hinsichtlich der Verteilung der Varianz auf Schulebene zeigt sich, dass lediglich 3.7% der Varianz in der emotionalen Erschöpfung auf Unterschiede zwischen den Schulen zurückzuführen ist. Dem gegenüber zeigt sich für die Kontextcharakteristiken (z. B. Unterstützung durch die Schulleitung: ICC = .22), dass ein substantieller Anteil der Varianz zwischen den Schulen liegt. Die Regressionsanalysen ergaben auf Schulebene keine signifikanten Prädiktoren, während auf individueller Ebene die subjektiv berichteten technischen Fähigkeiten, die Unterstützung durch die Schulleitung und die Zusammenarbeit im Kollegium mit einer geringeren Erschöpfung zusammenhingen. Lehrkräfte, die mehr gesundheitsbezogene Ängste und einen höheren Mehraufwand durch das Distanzlernen als andere Lehrkräfte im Kollegium angaben, berichteten hingegen eine höhere emotionale Erschöpfung.

Zusammenfassend zeigt sich, dass der individuellen Wahrnehmung der Lehrkräfte eine größere Bedeutung zukommt als der geteilten Wahrnehmung auf Schulebene. Einschränkend sind exemplarisch einige Limitationen wie die Verwendung eines querschnittlichen Designs zu berücksichtigen, wodurch keine kausalen Schlüsse gezogen werden können. Anhand der aktuellen Befunde wird ersichtlich, dass nur geringe Unterschiede zwischen den Kollegien im Niveau der emotionalen Erschöpfung verschiedener Schulen existieren. Das Erleben von emotionaler Erschöpfung steht dabei mit der individuellen Wahrnehmung des Schulkontextes in Zusammenhang; möglicherweise vor dem Hintergrund individueller Ressourcen und Coping-Strategien. Die Befunde werden anhand ihrer praktischen Implikationen diskutiert.

D-02: Evidenz und Kredibilität

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S06 S00 B32
Chair der Sitzung: Uwe Neugebauer

Evidenz und Unsicherheit: Nur der wahnsinnige Bildungswissenschaftler ist sich absolut sicher.

Uwe Neugebauer
BIBB, Deutschland

Im Gegensatz zur Implementationsforschung (z.B. Schrader & Hasselhorn, 2020) wird hier auf die individuelle Lehrkraft als Bildungswissenschaftler fokussiert: Wie werden von ihr bildungswissenschaftliche Evidenzen rezipiert, verarbeitet und bewertet, und wann führen Evidenzen zu Ablehnung, Einstellungsänderungen oder Unsicherheit? Mit der Erkenntnis zur *prinzipiellen* Unsicherheit des eigenen (Fach-)Wissens geht bei der Lehrkraft häufig eine Offenheit für neue Perspektiven und Lösungen einher. Im Zuge dieser Offenheit kann Evidenz relativ neutral verarbeitet werden: Personen mit einem stabilen beruflichen Selbstwert tendieren dazu, einstellungskonträre Evidenzen als Anlass zu nehmen, eine Einstellung zu revidieren.

Eine Lehrkraft mit geringem Selbstwert, wenig Offenheit und Ambiguitätstoleranz, die mit einstellungskonträren Evidenzen konfrontiert wird, verzerrt aus Selbstwertschutzgründen die Informationsverarbeitung und -bewertung über den Disconfirmation-Bias, und bei ihr findet kaum Einstellungsänderung statt.

Widersprüchliche Evidenzen erzeugen Unsicherheit, führen aber zu keiner Einstellungsveränderung. Auch unterscheiden sich Menschen in ihrem Kognitionsbedürfnis („*Need for Cognition*“, NFC-K, Beißert, Köhler, Rempel & Beierlein, 2014), ihrem Bedürfnis nach kognitiver Geschlossenheit und in ihrer Ambiguitätstoleranz (Herman, Stevens, Bird, Mendenhall & Oddou, 2010). Diese Faktoren moderieren Einstellungsänderungen (z.B. Haugtvedt & Petty, 1992).

Eine Einstellung wird verstärkt, wenn diese begründet werden soll, wie Studien zum Geschmack von Erdbeermarmelade (Wilson & Schooler, 1991) oder zu politischen Einstellungen (Strandberg, Sivén, Hall, Johansson & Pärnamets, 2018) zeigen. Hier soll untersucht werden, ob durch diese Verstärkung der eigenen Position die Veränderung durch einstellungskonträre Evidenz negativ beeinflusst wird.

Ableitungen und Hypothesen

Haupteffekt 1:

Wenn eine Person einstellungskonträre Evidenzen verarbeitet, wirkt der Diconfirmation-Bias: Evidenz wird dahingehend analysiert, was kritisierbar ist. Auch findet sie die Evidenz weniger valide und hilfreich als Personen, für die die Evidenz einstellungskonform ist.

Hypothese 1a: Einstellungskonforme Evidenzen werden als valider und hilfreicher eingeschätzt als einstellungskonträre Evidenzen.

Hypothese 1b: Einstellungskonforme Evidenzen werden kürzer betrachtet als einstellungskonträre Evidenzen.

Haupteffekt 2:

Durch eine Begründung der Einstellung wird die Einstellung einer Person rigider. Es findet weniger Einstellungsveränderung statt.

Hypothese 2: Wird eine Begründung VOR Darbietung der Evidenzen verlangt, ist die Einstellungsveränderung geringer als ohne Begründung.

Interaktionseffekt 1:

Die Einstellung wird nur verändert, wenn in sich konsistente einstellungskonträre Evidenzen vorgelegt werden UND ein stabiler Selbstwert sowie Ambiguitätstoleranz besteht.

Hypothese 3a: Bei einstellungskonträren Evidenzen korrelieren Einstellungsveränderung und Selbstwert (r = 0.35).

Hypothese 3b: Bei einstellungskonträren Evidenzen korrelieren Einstellungsveränderung und Ambiguitätstoleranz, dem Bedürfnis nach kognitiver Geschlossenheit sowie NFC-K (r = 0.35).

Methode

Die Befragung wird über LimeSurvey im November–Dezember 2022 umgesetzt. Als Stichprobe wurden Lehrkräfte gewählt, die über informelle Kontakte und social-media – Hinweise rekrutiert werden.

Vor Beginn des experimentellen Ablaufs wird Alter, Beruf und, ob Erfahrungen zu dem konkreten Thema bestehen, abgefragt. Daneben wird der NFC, der Selbstwert sowie die Ambiguitätstoleranz erhoben.

Der experimentelle Ablauf gliedert sich in drei Abschnitte:

In Abschnitt 1 wird bei der Lehrkraft (LK) mit 7 Fragen die Einstellung zu einem von zwei Fachthemen abgefragt (45 vs. 60 Minuten-Unterrichtseinheit; Lehrerraumprinzip vs. Klassenraumprinzip).

In Abschnitt 2 werden zwei empirische Evidenzen zum Thema präsentiert, jede Evidenz soll hinsichtlich Validität und, ob sie zur Einschätzung des Themas hilfreich war, bewertet werden.

Die Zuweisung der Evidenzen erfolgt über die Einstellungsabfrage in Abschnitt 1: 50% der LK erhalten einstellungskonträre, 25% einstellungskonforme sowie 25% mixed Evidenzen. Abschnitt 2 abschließend sollen jeweils zwei Multiple-Choice-Fragen zu den Evidenzen beantwortet werden.

In Abschnitt 3 wird erneut die Einstellung zum Thema abgefragt, als letzte Frage soll eingeschätzt werden, wie sicher / unsicher sich die Person in ihrer Einstellung sei.

Der Stichprobenumfang wurde mit G-Power (Faul, Erdfelder, Buchner & Lang, 2009) berechnet. Als Effektstärke wird für Hypothese 1 und 2 ein Effekt von d´ = 0.5 angenommen (Wilson & Schooler, 1991, Exp. 1), für Hypothese 3 ein Effekt von r = 0.35. Dies ergibt bei 5% a- und ß-Fehler 81 (Hypothese 3) bzw. 88 Personen.

Bound estimation of mean achievement scores with credible assumption about school-level nonresponse

Diego Cortes[1], Jeff Dominitz[2], Maximiliano Romero[3]
[1]IEA, Germany; [2]IEA, Germany; [3]ECONNorthwest

International large-scale assessments in the field of education (ILSAs) use probability samples to learn about the population-wide distribution of student achievement scores. The design commonly employed is a two-stage random sample, with schools sampled in a first stage and then a set of students or classes are selected across participating schools. This sampling strategy implies that unit

nonresponse during data collection can occur at either stage of the sampling process; that is, either at the school level or at the class/student level. In this study, we abstract from the latter source of nonresponse and focus attention solely on school-level nonresponse.

Unit nonresponse constitutes an analytical challenge in two different ways. First, it reduces the sample size and therefore the amount of information available for the estimation, acting upon the sampling distribution of the estimated parameters. Second, to derive point estimates of population parameters, nonresponse forces researchers to make distributional assumptions about the unobserved measures of nonresponding units. In this research project we focus attention in the latter.

We study the analytical challenge that survey nonresponse imposes to ILSA studies. Its contribution to the field of survey methodology in education is threefold. First, it studies how different assumptions about the unobserved measures in nonresponding schools can impact the inference reached about the distribution of achievement scores in a student population. Second, it provides a fresh perspective on how to measure the ambiguity in the inference caused by nonresponding units. Finally, it introduces the partial identification framework to ILSA field of work and uses it as an analytical tool to study survey nonresponse.

Reports in ILSA studies assume that school-level nonresponse is ignorable within a study (sub)population. This is a convenient assumption, as it allows point-identification of the parameter of interest. However, it is a strong assumption, as it maintains that measures in nonresponding schools follow the exact distribution as the measures in responding schools. We study how inference can change if we replace this assumption with weaker and more credible assumptions on survey nonresponse.

The partial identification framework (Manski, 1995, 2003) provides clarity about what can be learned about a probability distribution based on the data generating process if we were to make minimal or no assumptions about survey nonresponse. Let each member of a student population be characterized by the duplet (y,z). Where y denotes the student achievement scores, and z is an indicator specifying whether a school participates in a survey. A central goal in ILSAs is to identify the probability distribution of achievement scores in a population P(y).

Imposing minimal assumptions about nonresponding units, we construct bounds of the mean achievement score in a study population E(y). Using the Law of Iterated Expectations, the identification region of E(y) can be expressed as:

H{E(y)}={E(y|z=1)P(z=1)+y_0 P(z=0),E(y|z=1)P(z=1)+y_1 P(z=0)}

Where, P(z=1) and P(z=0) denote the probability of participation and not participation, respectively. H{E(y)} can be estimated using different assumptions about y_0 and y_1.

We use data from all populations participating in the International Computer and Information Literacy Study (ICILS) to estimate the bounds in Equation 1. The empirical evidence gathered by ICILS reveals E(y|z=1), P(z=1), and P(z=0), but reveals nothing about the distribution of scores for students in not participating schools. To estimate the bounds of H{E(y)} in Equation 1, we therefore must assume something about y_0 and y_1. In this project, to estimate the lower (upper) bound of H{E(y)} we assume that P(y | z=0) is degenerate at y_0 (y_1). We present different bounds H{E(y)} under different values of y_0 and y_1.

Bildungsmythen mit Refutation Texts widerlegen: Machen Leseziele einen Unterschied?

Jana Asberger, Eva Thomm, Johannes Bauer
Universität Erfurt, Deutschland

Ein erfolgsversprechender Ansatz, um Bildungsmythen zu widerlegen, sind Refutation Texts (Zengilowski et al., 2021). Sie provozieren einen kognitiven Konflikt, indem sie eine Fehlannahme explizit benennen und mit einer wissenschaftlichen Erklärung korrigieren (Tippett, 2010). Bisher gibt es jedoch noch wenige Erkenntnisse darüber, welche Faktoren die Wirkung von Refutation Texts begünstigen (Zengilowski et al., 2021). Ein relevanter Einflussfaktor könnten Leseziele sein, da sie Leseprozesse und -ergebnisse beeinflussen (Trevors & Muis, 2015). In der Leseforschung werden häufig zwei bestimmte Leseziele aus dem Alltag von Studierenden gegenübergestellt: Lesen zum Lernen und Lesen zur Unterhaltung (z.B. Linderholm & van den Broek, 2002). Mehrere Experimente zeigten, dass Studierende, bei denen ein lernorientiertes Leseziel induziert wurde, elaborierter lasen und sich besser an den Textinhalt erinnerten als bei einem unterhaltungsorientierten Leseziel (McCrudden & Schraw, 2007). Lernorientierte Leseziele könnten folglich die Entstehung des gewünschten kognitiven Konflikts zwischen Fehlannahmen und neuen Informationen begünstigen und damit auch den Refutation Effekt.

Die vorliegende Studie untersuchte vor diesem Hintergrund, ob Leseziele die Wirksamkeit eines Refutation Texts zu einem Bildungsmythos (Bedeutsamkeit der Hemisphärendominanz) moderieren. Wir erwarteten, dass Lehramtsstudierende, die einen Refutation Text lesen, ihre Fehlannahme eher ändern würden, als Lehramtsstudierende, die einen traditionellen Text lesen (H1a). Zudem nahmen wir an, dass dieser Effekt stärker ausfällt, wenn sie einen Text lesen, um eine Erklärung zu erhalten als wenn sie zur Unterhaltung lesen (H1b). Da Konzeptwechsel damit beginnen kann, dass man sich seiner Überzeugung unsicherer wird (Taylor & Kowalski, 2004), nahmen wir weiterhin an: Lehramtsstudierende, die an den Mythos glauben, äußern eine geringere Sicherheit in ihrem Fehlkonzept, zeigen also ein angemesseneres Sicherheitsurteil, wenn sie einen Refutation Text als einen traditionellen Text lesen (H2a). Dieser Effekt ist stärker, wenn sie einen Text lesen, um eine Erklärung zu erhalten als wenn sie zur Unterhaltung lesen (H2b).

Das Online-Experiment (N=168) folgte einem 2x2x2-mixed-Design mit den Zwischensubjektfaktoren Textart (Refutation vs. traditionell) sowie Leseziel (Erklärung vs. Unterhaltung) und dem Innersubjektfaktor Zeit (vor vs. nach Textlektüre). Die Teilnehmer*innen wählten zu beiden Messzeitpunkten aus, welcher von zwei Aussagen zur Hemisphärendominanz sie eher zustimmten und schätzten auf einer Skala von 50% bis 100% ein, wie sicher sie sich in ihrer Zustimmung waren. Anschließend wurden die Teilnehmer*innen instruiert einen Text zu lesen, um entweder eine Erklärung zu erhalten oder um ihn zu Unterhaltungszwecken zu lesen. Bei den Texten wurde manipuliert, ob es sich um einen traditionellen Text zum Thema Hemisphärendominanz oder um einen Refutation Text handelte. Zur Überprüfung der Hypothesen führten wir eine binäre logistische Regression (H1a,b) und eine Varianzanalyse (H2a,b) durch.

Die Ergebnisse zeigten, dass Teilnehmer*innen, die einen Refutation Text lasen, eher ihre Überzeugung änderten als Teilnehmer*innen, die einen traditionellen Text lasen (H1a).

Die Wirkung des Refutation Effekts war, anders als erwartet, unabhängig vom Leseziel (H1b). Allerdings gab es einen signifikanten Haupteffekt des Leseziels. Teilnehmer*innen änderten eher ihr Fehlkonzept, wenn sie zur Unterhaltung lasen als wenn sie den Text lesen sollten, um eine Erklärung zu erhalten.

Die ANOVA ergab, dass Teilnehmer*innen, die den Refutation Text lasen, entweder ihr Fehlkonzept änderten oder sich danach unsicherer in ihrer Überzeugung waren (H2a). Der traditionelle Text verunsicherte die Teilnehmer*innen nicht bezüglich ihres Fehlkonzeptes.

Das Leseziel moderierte wider Erwarten nicht den Refutation Effekt auf das Sicherheitsurteil (H2b). Die Interaktion von Leseziel und Zeit war jedoch signifikant. Teilnehmer*innen, die zur Unterhaltung lasen, zeigten anschließend ein angemesseneres Sicherheitsurteil. Teilnehmer*innen, die den Text lesen sollten, um eine Erklärung zu erhalten, ließen sich in ihrem Fehlkonzept nicht verunsichern.

Die Ergebnisse stützen die Wirksamkeit von Refutation Texts. Entgegen der Erwartungen zeigte sich nicht, dass Leseziele den Refutation Effekt moderieren. Das Befundmuster sollte auch mit anderen Bildungsthemen überprüft werden. Implikationen für die Gestaltung wissenschaftlicher Texte und Instruktionen ihrer Nutzung im Lehramtsstudium werden diskutiert.

Optimierung der deutschen Stratifizierung in PISA: eine Simulationsstudie

Julia Mang[1], Sabine Meinck[2], Helmut Küchenhoff[3]
[1]TU München/ZIB, München, Deutschland; [2]International Association for the Evaluation of Educational Achievement (IEA), Hamburg, Deutschland; [3]Ludwig-Maximilians-Universität München, Institut für Statistik, München, Deutschland

Stratifizierung ist ein wichtiges Merkmal des komplexen Stichproben-Designs, welches in Large-Scale Assessments (LSA) wie dem Programme for International Student Assessment (PISA) angewandt wird (z.B. OECD, 2020). Wenn Kenngrößen, die für die Stratifizierung genutzt werden, mit den Zielgrößen, d.h. den Kompetenzen der 15-jährigen Schüler*innen korreliert sind, kann die Effizienz der Studie verbessert werden (Jaeger, 1984). Dabei erhöht sich die Stichproben-Schätzgenauigkeit für diese Variablen, während der Gesamtstichprobenumfang konstant bleibt (Cochran, 1977; Meinck & Vandenplas, 2021). Darüber hinaus kann eine Stratifizierung sicherstellen, dass alle Teile einer Population in einem vorher festgelegten Verhältnis in die Stichprobe aufgenommen werden. Dabei wird zwischen expliziter und impliziter Stratifizierung unterschieden. Explizite Stratifizierung bedeutet, dass alle teilnahmeberechtigten Schulen (d.h., solche mit 15-jährigen Schüler*innen) in Untergruppen eingeteilt werden, wobei alle Schulen, die zu einer Untergruppe gehören, als eine separate Gruppe behandelt werden. Implizite Stratifizierung beinhaltet, dass diese getrennten Gruppen nach bestimmten Merkmalen sortiert werden (Meinck, 2020).
In dieser Arbeit wird auf Grundlage einer Simulationsstudie ermittelt, welche Stratifizierungsvariante für PISA in Deutschland optimal ist. Hierbei wurden die durch das internationale PISA-Komitee vorgegebenen Rahmenbedingungen, die verfügbaren Informationen über Schulen und ihre Schüler*innen in statistischen Ämtern des Bundes und der Länder und die spezifischen Merkmale des deutschen Bildungssystems berücksichtigt (OECD, 2020). Wir untersuchen sieben verschiedene Stratifizierungsvarianten. Diese basieren auf Szenarien, welche in früheren LSAs in Deutschland verwendet wurden, sowie auf theoretischen Überlegungen für zukünftige Implementierungen und Verbesserungen. Die ausgewählten Ansätze wurden zusätzlich mit einem Szenario ohne

Stratifizierung und einer synthetischen optimalen Stratifizierung verglichen, um ein umfassendes Bild zu erhalten.

Die Simulationsstudie zeigt, dass die derzeit angewandte Stratifizierung zufriedenstellende Ergebnisse hinsichtlich der Stichprobengenauigkeit erzielt. Das aktuelle Design für PISA basiert dabei auf den Bundesländern als explizite und Schulformen als implizite Stratifizierung in Deutschland (Mang et al., 2020). Ein schwerwiegendes Problem dieses Ansatzes besteht jedoch darin, dass er zu kleine Stratifizierungsgruppen ergeben kann, was bei der Schätzung der Stichprobenvarianz in früheren Zyklen in PISA zu Problemen geführt hat. Daher wurden alternative Szenarien für die Stratifizierung in dieser Studie betrachtet, die - bei gleichzeitigem Lösen dieses Problems - auch zu besseren Ergebnissen führten, d. h. zu kleineren Standardfehlern (SE) für die geschätzten Schülerleistungen in Mathematik, Naturwissenschaften und Lesen.

Als Ergebnis dieser Studie empfehlen wir, in künftigen PISA-Zyklen drei verschiedene Stratifizierungsvarianten für Deutschland in Betracht zu ziehen, die (i) die Stichprobeneffizienz verbessern und gleichzeitig den Stichprobenumfang konstant halten, (ii) einem soliden methodischen Ansatz folgen und (iii) eine konservative und vorsichtige Änderung darstellen, indem die Struktur Deutschlands mit seinen Bundesländern und Schultypen erhalten bleibt. Eine zu empfehlende Stratifizierungsvariante verwendet eine Gruppierung der Bundesländer in drei Kategorien (Stadtstaaten, alte und neue Bundesländer). Da diese Gruppierung im Vergleich zur bisherigen, verwendeten Stratifizierung die föderale Länderstruktur Deutschlands beibehält, kann sie eine gute Alternative für das Stratifizierungsdesign der kommenden PISA-Zyklen darstellen, da sie eine konservative und vorsichtige Änderung impliziert. Durch eine zusätzliche implizite Schichtung nach Bundesländern kann das Problem der unvorhersehbaren Stichprobengrößen pro Bundesland gelöst werden, da dieses Verfahren zu einer nahezu perfekten proportionalen Verteilung der Stichprobe auf alle Schichten führt. Zwei unserer Empfehlungen nutzen zusätzlich die Schulformen als expliziten Schichtung, was eine hohe Stichprobengenauigkeit gewährleistet, da die Schulform sehr eng mit der durchschnittlichen Kompetenz der Schüler*innen zusammenhängt. Insgesamt sind wir der Ansicht, dass die Stratifizierung mit den Schulformen als explizite und den Bundesländern als implizite Stratifizierung alle Anforderungen an ein Stratifizierungsdesign in Deutschland erfüllt und daher für künftige PISA-Zyklen uneingeschränkt empfohlen werden kann.

D-03: Wissenschaftskommunikation

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S06 S00 B41
Chair der Sitzung: Mark Jonas

Am besten einfach und wissenschaftlich? Der Einfluss von Einfachheit und Wissenschaftlichkeit auf Vertrauen und Zusammenhänge mit Leser:innenmerkmalen

Mark Jonas, Martin Kerwer, Anita Chasiotis, Tom Rosman
Leibniz-Institut für Psychologie (ZPID), Germany

Theoretischer Hintergrund. Vor dem Hintergrund aktueller gesellschaftlicher Entwicklungen wie Fake News (Tandoc, 2019), Verschwörungstheorien oder Wissenschaftsleugnung (Lewandowsky et al., 2017) stellt sich zunehmend die Frage, welche Faktoren das Vertrauen von Leser:innen in Wissenschaftler:innen und Wissenschaft beeinflussen. Bisherige Forschung liefert Evidenz dafür, dass im Rahmen wissenschaftlicher Texte sowohl die wahrgenommene Einfachheit ("Easiness-Effekt", Scharrer et al., 2012, 2017) als auch Wissenschaftlichkeit ("Scientificness-Effekt", Bromme et al., 2015; Thomm & Bromme, 2012) von Texten Vertrauen positiv beeinflussen können. Unklar ist allerdings, in welchem Ausmaß Wechselwirkungen zwischen Einfachheit und Wissenschaftlichkeit Vertrauenseinschätzungen beeinflussen und welche individuellen Merkmale von Leser:innen eine Rolle spielen. Es erscheint z.B. plausibel, dass individuelle epistemische Überzeugungen (epistemic justification beliefs, Ferguson et al., 2012) oder das Bedürfnis nach kognitiver Geschlossenheit (need for cognitive closure, NCC; siehe Kruglanski et al., 2010) von Leser:innen mit den oben genannten Effekten interagieren könnten.

Fragestellung. Im Vortrag soll, basierend auf Ergebnissen einer experimentellen Online-Studie, die Frage beantwortet werden, inwiefern die Einfachheit und Wissenschaftlichkeit von Texten sowie deren Interaktion Vertrauen prädizieren können. Zudem sollen exploratorische Befunde zum Einfluss individueller epistemischer Überzeugungen und NCC auf den Easiness- und Scientificness-Effekt vorgestellt werden.

Methode. In einer präregistrierten Online-Studie mit einem messwiederholten 2(Einfachheit: hoch vs. niedrig) x 2(Wissenschaftlichkeit: hoch vs. niedrig) within-Design wurden N = 1467 nach Geschlecht, Alter und Bildungsgrad stratifizierte Leser:innen aus der Allgemeinbevölkerung untersucht. Den Versuchspersonen wurden in zufälliger Reihenfolge vier fiktive Kurzzusammenfassungen auf Basis tatsächlich veröffentlichter psychologischer Fachrtikel präsentiert. Jeder Text deckte randomisiert eine der vier möglichen Kombinationen aus Einfachheit und Wissenschaftlichkeit ab. Einfachheit wurde über Satzlänge, Nebensätze, Wortlänge und Strukturierung durch Absätze und Stichpunkte manipuliert, während Wissenschaftlichkeit durch das Ausmaß an neutraler oder wertender Sprache, an Referenzen und Statistik sowie durch den Detailgrad von Methodenbeschreibungen manipuliert wurde. Leser:innen gaben für jeden Text eine Einschätzung ihres Vertrauens in die Textaussagen auf einer 8-stufigen Likert-Skala an. Darüber hinaus wurde ihr Vertrauen in die jeweiligen Wissenschaftlerinnen mittels des Muenster Epistemic Trustworthiness Inventory (METI, Hendriks et al., 2015) erhoben. Am Ende der Studie bearbeiten Leser:innen zudem einen auf Klopp und Stark (2016) basierenden Fragebogen zu epistemischen Überzeugungen sowie einen Fragebogen zu individuellem NCC (Schlink & Walther, 2007).

Ergebnisse. Der Einfluss von Einfachheit, Wissenschaftlichkeit und ihrer Interaktion auf Vertrauen wurde mittels Mixed-Model-Analysen ausgewertet. Eine höhere Wissenschaftlichkeit von Texten

sagte sowohl für die Textaussagen als auch für die Wissenschaftlerinnen signifikant höheres Vertrauen vorher (β = .142 – .512, SE =.027 – .033, p < .001, R^2 = .003 – .037). Dieser Zusammenhang zeigte sich allerdings weder für Einfachheit (alle ps >.343) noch für die Interaktion zwischen Einfachheit und Wissenschaftlichkeit (alle ps >.285). Mit Blick auf die epistemischen Überzeugungen von Leser:innen zeigten sich Interaktionen zwischen Wissenschaftlichkeit und personal justification beliefs (β = -.305 – -.156 , SE = .019 – .023, p <.001, R^2 = .007 – .022) sowie justification by multiple sources (β = .079 – .145, SE = .019 – .024, p <.001, R^2 = .002 – .006). Wenn Leser:innen wissenschaftliches Wissen als subjektiv und vorläufig einschätzten (personal justification), so ergaben sich geringere Vertrauensunterschiede zwischen Zusammenfassungen hoher und niedriger Wissenschaftlichkeit. Ein umgekehrter Effekt fand sich für Leser:innen mit einer starken Überzeugung, dass sich Wissen nur durch das Vergleichen und Synthetisieren von Befunden rechtfertigen lässt (justification by multiple sources): Hier ergaben sich mit zunehmender Überzeugung größere Vertrauensunterschiede zwischen Zusammenfassungen hoher und niedriger Wissenschaftlichkeit.

Diskussion. Die vorliegenden Befunde weisen auf die Bedeutsamkeit eines wissenschaftlichen Kommunikationsstils zur Förderung des Vertrauens von Leser:innen aus der Allgemeinbevölkerung hin. Zudem liefern sie Anhaltspunkte dafür, dass Texte mit hoher Wissenschaftlichkeit insbesondere bei Personen mit “wissenschaftsnäheren” epistemischen Überzeugungen vertrauensbildend sind. Zum Abschluss des Vortrags sollen das Ausbleiben des Easiness-Effekts und mögliche Implikationen der Studie für die (schriftliche) Wissenschaftskommunikation diskutiert werden.

Belief updating and trust: Der Einfluss von Wissenschaftsvertrauen auf die Veränderung themenspezifischer Überzeugungen im Rahmen einer Konfrontation mit wissenschaftlicher Evidenz

Tom Rosman, Sianna Grösser
Leibniz-Institut für Psychologie (ZPID)

Theoretischer Hintergrund

Evidenzbasiertes Denken und Handeln setzt eine positive Bewertung der jeweiligen Evidenz (z. B. wissenschaftlicher Befunde) voraus. Wissenschaftlichen Laien fehlt es jedoch meist an Wissen und Fertigkeiten, um wissenschaftliche Befunde adäquat einzuschätzen. Aus diesem Grund bewerten sie die Evidenz oftmals nicht direkt, sondern beurteilen die Vertrauenswürdigkeit der jeweiligen Quelle, also beispielsweise der Wissenschaftlerin, die über ihre Forschungsergebnisse berichtet (Bromme et al., 2010). Somit kommt Wissenschaftsvertrauen immer dann, wenn Individuen mit wissenschaftlicher Evidenz konfrontiert sind, eine zentrale Rolle bei der Veränderung themenspezifischer Überzeugungen (sog. belief updating; Hogarth & Einhorn, 1992; Pilditch et al., 2020) zu. Trotz großer Plausibilität dieser Annahme und einem vermehrten Aufkommen entsprechender theoretischer Argumente (z. B. Druckman & McGrath, 2019) gibt es jedoch vergleichsweise wenig empirische Studien dazu.

Fragestellung

Der vorliegende Beitrag untersucht, inwiefern eine Konfrontation mit wissenschaftlicher Evidenz ein belief updating im Sinne einer Veränderung themenspezifischer Überzeugungen zum Thema

„Akupunktur bei Rückenschmerzen“ evozieren kann, und welche Rolle Wissenschaftsvertrauen dabei spielt. Es wird angenommen, dass belief updating mit höherem Wissenschaftsvertrauen stärker ausgeprägt ist.

Methode

Zur Untersuchung der Fragestellung wurden Daten einer experimentellen Studie mit insgesamt *N*=985 Personen herangezogen (deutschsprachige Allgemeinbevölkerung von 18-65 Jahren; Altersdurchschnitt: *M*=42.42; *SD*=13.47; 48.8% weiblich). Die Studie war als 2x3 within-between-Design angelegt. Den Experimentalfaktor stellte eine Leseaufgabe dar, in deren Rahmen die Versuchspersonen 12 Kurzzusammenfassungen fiktiver wissenschaftlicher Studien lasen. In der Bedingung „pro Akupunktur“ enthielt die Aufgabe acht Studien, die eine Überlegenheit von Akupunktur (gegenüber Massagen) nahelegten (und vier thematisch irrelevante Kontrollstudien). In der Bedingung „contra Akupunktur“ indizierten die acht Studien eine Unterlegenheit von Akupunktur. In der Kontrollgruppe (Bedingung „divergierende Evidenz“) indizierten schließlich vier Studien eine Überlegenheit und vier Studien eine Unterlegenheit. Den within-Faktor stellte die Prä-Post-Messung von akupunkturbezogenen Überzeugungen bzw. beliefs dar, welche mit einer eigens konstruierten 8-Item-Skala gemessen wurden (Beispielitem: „Mit Akupunktur lassen sich Rückenschmerzen deutlich reduzieren“; sechsstufiges Antwortformat; α=.78 [prä] bzw. α=.85 [post]). Wissenschaftsvertrauen wurde schließlich mithilfe einer Übersetzung der 5-Item-Skala von Nisbet et al. (2015) erhoben (Beispielitem: „Erkenntnisse von Wissenschaftler/innen sind vertrauenswürdig.“; sechsstufiges Antwortformat; α=.79).

Ergebnisse und Diskussion

Zur Prüfung, inwiefern eine Konfrontation mit wissenschaftlicher Evidenz eine Veränderung themenspezifischer Überzeugungen evozieren kann, wurde eine zweifaktorielle ANOVA mit Messwiederholung durchgeführt (between-Faktor: experimentelle Bedingung, within-Faktor: akupunkturbezogene Überzeugungen). Es zeigte sich eine signifikante within-between-Interaktion (F[2,981]=178.76; p<.001). Erwartungsgemäß gab es einen Anstieg der akupunkturbezogenen Überzeugungen in der „pro Akupunktur“-Gruppe (im Sinne einer positiveren Bewertung von Akupunktur), ein Absinken in der „contra Akupunktur“-Gruppe, und nur wenig Veränderung in der Kontrollgruppe. Tukey HSD post hoc-Tests zeigten signifikante Unterschiede zwischen den jeweiligen Gruppen an (alle p<.01).

Zur Überprüfung der Hauptfragestellung, also der Interaktionen der experimentellen Manipulation mit Wissenschaftsvertrauen, wurde das SPSS-Makro PROCESS 4.1 herangezogen (Hayes, 2018). Abhängige Variable war hierbei die Differenz zwischen prä- und post-Messung akupunkturbezogener Überzeugungen (Montoya, 2019), Vertrauen fungierte als (zentrierte) unabhängige Variable, und experimentelle Bedingung als Moderator (dummykodiert; Referenzgruppe: Kontrollgruppe). Es zeigten sich signifikante Interaktionen zwischen Wissenschaftsvertrauen und der „pro Akupunktur“-Dummyvariable (B=-0.15; p<.01; 95% CI [-0.27, -0.04]) sowie der „contra Akupunktur“-Dummyvariable (B=0.16; p<.01; 95% CI [0.05, 0.28]). Auf Gruppenebene betrachtet zeigte sich in der „pro Akupunktur“-Bedingung ein positiver Effekt von Vertrauen auf die Überzeugungsänderung hin zu

einer positiveren Bewertung von Akupunktur (p<.01). In der „contra-Akupunktur"-Bedingung war dieser Effekt erwartungsgemäß negativ (p<.001), und in der Kontrollgruppe erwartungsgemäß nicht signifikant (p=.32). Die Effekte der experimentellen Manipulationen in den beiden Experimentalgruppen wurden mit steigendem Wissenschaftsvertrauen somit zunehmend stärker.

Die Befunde belegen die zentrale Rolle von Wissenschaftsvertrauen als Einflussfaktor beim belief updating. Ungeklärt ist allerdings, ob Personen mit höherem Wissenschaftsvertrauen aufgrund dessen womöglich auch leichter manipulierbar sind durch fragwürdige wissenschaftliche Evidenz oder Pseudowissenschaft. Diese Fragestellung soll in einer Folgestudie geklärt werden. Implikationen der Befunde werden diskutiert.

Wie bewerten Laien die Kommunikation von wissenschaftlicher Unsicherheit mit Hinblick auf die Vertrauenswürdigkeit von Wissenschaftler*innen?

Inse Janssen, Regina Jucks
Westfälische Wilhelms-Universität Münster, Germany

Theoretischer Hintergrund

Unsicherheit als fundamentale Eigenschaft von Wissenschaft zu verstehen und gleichzeitig wissenschaftliche Erkenntnis nicht ihre Verlässlichkeit abzusprechen kann herausfordernd für Laien sein (Kampourakis & McCain, 2020). Zuletzt ist dies auch während der Coronapandemie deutlich geworden: Einerseits ist die Offenlegung von Unsicherheit für eine informierte Entscheidungsgrundlage besonders dann unabdingbar, wenn wissenschaftliche Informationen in einem direkten Zusammenhang mit politischen und persönlichen Entscheidungsprozessen stehen (Blastland et al., 2020). Andererseits könnte wissenschaftliche Unsicherheit auch missverstanden werden, Erkenntnisse deshalb als nicht handlungsleitend eingestuft werden oder Wissenschaft gänzlich als Informationsquelle abgewertet werden. Dies sind aber vor allem vermeintliche Folgen von Unsicherheitskommunikation, die Wissenschaftler*innen selbst befürchten (Post, 2016; Post & Maier, 2016).

Empirische Studien zeigen hingegen, dass Laien in den meisten Fällen wissenschaftliche Unsicherheit richtig einordnen können und nicht als Schwäche verstehen (Gustafson & Rice, 2020). Vermutlich könnte die transparente Kommunikation von Unsicherheit sogar dazu beitragen, dass Laien Wissenschaftler*innen als vertrauenswürdige Informationsquelle, d. h. als kompetent, integer und wohlwollend wahrnehmen. Schließlich signalisiert die Kommunikation von Unsicherheit ehrliche und objektive Kommunikationsabsichten von Wissenschaftler*innen (Steijaert et al., 2020). Besonders wenn Wissenschaftler*innen sich in öffentlichen Debatten zu Wort melden, und die Grenzen zwischen Wissenschafts- und politischer Kommunikation verschwimmen, könnte die Offenlegung von Unsicherheit außerdem verhindern, dass Laien Wissenschaftler*innen als persuasive Akteure wahrnehmen, die nur bestimmte Interessen durchsetzen möchten (Abdool Karim, 2022).

Forschungsfrage

In einem 2x2 Vignettenexperiment sind wir diesen Annahmen weiter nachgegangen und haben folgende Forschungsfrage untersucht: Wie wirkt sich die Kommunikation von wissenschaftlicher Unsicherheit auf die wahrgenommene Vertrauenswürdigkeit von Wissenschaftler*innen aus, wenn sie

auch ihre Absicht äußern sich für politische Handlungsempfehlungen einzusetzen (vs. nur die Absicht äußern über wissenschaftliche Erkenntnisse zu informieren)?

Methode

In der im Mai 2022 durchgeführten Online-Studie wurden 503 Versuchsteilnehmende gebeten ein fiktives, aber evidenzbasiertes Interview mit einer Wissenschaftlerin über den Nutzen weiterer und angepasster Impfungen gegen Covid-19 zu lesen. In dem Interview wurde erstens variiert, ob die Wissenschaftlerin entweder explizit wissenschaftliche Unsicherheit kommuniziert (z. B. Nennung von Limitationen oder Hinweis auf Vorläufigkeit) oder nicht. Darüber hinaus wurde zweitens in dem Interviewtext manipuliert, ob die Wissenschaftlerin entweder ihre Absicht darlegte lediglich über den aktuellen wissenschaftlichen Stand zu informieren, oder ihre Absicht ausdrückte, für politische Empfehlungen einzutreten. Anschließend bewerteten die Versuchsteilnehmenden die Vertrauenswürdigkeit der Wissenschaftlerin auf den drei Subskalen Expertise, Integrität und Wohlwollen (Hendriks et al., 2015).

Ergebnisse

Die Ergebnisse zeigen, dass die Versuchsteilnehmenden die Wissenschaftlerin als weniger vertrauenswürdig (d. h. weniger kompetent, integrer und wohlwollender) wahrgenommen haben, wenn sie Unsicherheit *nicht* explizit kommunizierte. Die Offenlegung der Absicht sich für politische Handlungsempfehlungen einzusetzen (vs. lediglich über Forschungsergebnisse informieren zu wollen) hatte keinen negativen Einfluss auf die wahrgenommene Expertise oder Integrität der Wissenschaftlerin. Allerdings führte die Offenlegung sich für politische Handlungsempfehlungen einzusetzen zu einer niedrigeren Bewertung des Wohlwollens – und das unabhängig davon, ob die Wissenschaftlerin explizit wissenschaftliche Unsicherheit explizit kommunizierte oder nicht.

Implikationen

Die Ergebnisse der Studie zeigen, dass Laien nicht etwa die explizite Kommunikation von Unsicherheit negativ bewerten, sondern das Gegenteil ist der Fall: das Fehlen der expliziten Kommunikation von Unsicherheit hat die zugeschriebene Vertrauenswürdigkeit einer Informationsquelle negativ beeinflusst. Gleichzeitig wird das Offenlegen von wissenschaftlicher Unsicherheit nicht als Freifahrtschein für persuasive Kommunikation eingeordnet. Dies deutet darauf hin, dass Laien die Kommunikation von wissenschaftlicher Unsicherheit als wichtige Information interpretieren, aber z. B. auch die dargelegten Kommunikationsabsichten einer Quelle berücksichtigen. Weitere Implikationen für den Umgang mit wissenschaftlicher Unsicherheit in Bildungskontexten werden diskutiert.

Das Projekt PLan Psy - Evidenzbasierte Entwicklung eines Angebots zur laiengerechten Kommunikation psychologischer Forschungsergebnisse

Martin Kerwer[1], Marlene Stoll[1,2], Gesa Benz[1], Mark Jonas[1], Pawel Nuwaltzew[1], Anita Chasiotis[1]
[1]Leibniz-Institut für Psychologie (ZPID), Deutschland; [2]Leibniz-Institut für Resilienzforschung (LIR), Deutschland

Theoretischer Hintergrund. Kognitive Arbeitsteilung ist ein integraler Teil unserer modernen Wissensgesellschaft (Bromme & Goldman, 2014). Hieraus resultiert die Notwendigkeit, Wissensbestände allgemeinverständlich aufzubereiten. Die Biomedizin verwendet hierzu bereits seit

mehreren Jahrzehnten erfolgreich Plain Language Summaries (PLS; allgemeinverständliche Kurzzusammenfassungen), wobei Cochrane-Reviews und PLS zugehöriger Metaanalysen (Santesso et al., 2015) als Goldstandard laiengerechter Kommunikation medizinischer Evidenz gelten. Auch in der empirischen Bildungsforschung erfreuen sich Clearing House Ansätze immer stärkerer Verbreitung und können zur Empiriebasierung pädagogischer Praxis beitragen (Seidel, Mok, Hetmanek & Knogler, 2017). Dieser angestrebten Empiriebasierung der Praxis steht allerdings gegenüber, dass Wissenschaftskommunikation selbst erst in jüngster Zeit und wenig systematisch empirisch beforscht wird, und die Praxis der Wissenschaftskommunikation damit unzureichend empirisch fundiert ist. Am Leibniz-Institut für Psychologie (ZPID) wird derzeit im Projekt PLan Psy ein Angebot entwickelt, welches metaanalytische psychologische Evidenz für die Allgemeinheit aufbereitet. Um die Lücke einer mangelnden empirischen Fundierung von Wissenschaftskommunikation zu adressieren, findet die Entwicklung dieses Angebots streng evidenzbasiert statt.

Fragestellung. Der Vortrag illustriert, anhand der Entwicklung der Richtlinie zum Verfassen von “KLARtexten” zu psychologischen Übersichtsarbeiten und Metaanalysen innerhalb des Projekts PLan Psy, wie Wissenschaftskommunikation empirisch fundiert gelingen kann.

Methode. Zur Entwicklung der Richtlinie wurde zunächst ein disziplinübergreifender systematischer Review zur theoretischen und empirischen Beforschung sowie zu existierenden Richtlinien zum Verfassen von PLS durchgeführt. Forschungsfragen und Forschungsbedarfe als Ausgangspunkte experimenteller Studien sowie bereits ausreichend fundierte Kriterien wurden hieraus abgeleitet. In drei präregistrierten quantitativen Studien wurde jeweils ein experimentelles Between-Person-Design innerhalb von Onlinestudien mit Stichproben von etwa 2000 Personen aus der Allgemeinbevölkerung - nach Alter, Geschlecht und Bildungshintergrund stratifiziert - umgesetzt (n = 2288, n = 2211, n = 2451). Teilnehmende lasen jeweils Zusammenfassungen zu zwei psychologischen Metaanalysen. Variiert wurden PLS-Charakteristika zu im Review identifizierten Evidenzlücken - beispielsweise Vorgehensweisen zum Umgang mit Fachbegriffen oder die Anzahl der berichteten Effekte. Eine vierte Experimentalstudie ist derzeit in Arbeit, wird aber zum Zeitpunkt des Vortrags abgeschlossen sein. Zur weiteren Untersuchung der Nutzer:innenwahrnehmung erster PLS-Prototypen wurde eine präregistrierte Fokusgruppenstudie mit interessierten Nicht-Psycholog:innen, Praktiker:innen (z.B. Psychotherapeut:innen) und Wissenschaftsjournalist:innen durchgeführt. Ergänzt wurde diese empirische Forschung durch Evaluationsworkshops mit Expert:innen aus den Bereichen Wissenschaftskommunikation, Metaanalysen und Forschungsinfrastruktur zu ersten Richtlinien-Versionen.

Ergebnisse. Innerhalb des Vortrags werden wesentliche Befunde der empirischen Arbeiten des Projekts in aggregierter Form skizziert. Die drei experimentellen Studien zeigten unter anderem, dass richtlinienkonforme PLS im Vergleich zu wissenschaftlichen Abstracts als verständlicher wahrgenommen werden (η^2 = .040, p<.001) und den Wissenserwerb steigern (OR = 1.33, p<.05). Hierbei wiesen KLARtexte, die Fachbegriffe durch Alltagsbegriffe ersetzen - statt sie beizubehalten und ausführlich zu umschreiben - einen besseren Wissenserwerb (OR = 1.54, p<.05) auf, und KLARtexte, die statistische Effektgrößen nur einordneten - aber nicht erläuterten zeigten eher eine

bessere Nutzer:innenerfahrung und Wissenserwerb. Erläuterungen zur Methode "Metaanalyse" (OR = 1.29, p<.01) und dem Problem des “Publication Bias” (OR = 2.19, p<.01) steigerten zudem das konzeptuelle Wissen zu diesen Themen. Auch die Relevanz von Informationen zu Interessenkonflikten konnte belegt werden (alle η^2 > .002, alle p<.05). Basierend auf den Rückmeldungen der ersten drei Studien wurde eine Pilotversion der Richtlinie erstellt. Die Ergebnisse der vierten Experimentalstudie, der Fokusgruppenstudie sowie eines zweiten Evaluationsworkshops der Richtlinie, fließen derzeit in die Weiterentwicklung der Richtlinie ein. Zum Zeitpunkt des Vortrags wird damit eine zweite überarbeitete Richtlinienversion vorliegen.

Diskussion. Zum Abschluss des Vortrags werden die wissenschaftliche Relevanz sowie praktische Implikationen der empirischen Projektstudien reflektiert. Dies umschließt aktuelle Schritte zur Überführung der Projektergebnisse in ein dauerhaftes, öffentlich zugängliches Informationsangebot am ZPID, welches zum Zeitpunkt des Vortrags im Produktivbetrieb sein wird. Durch die Fachgemeinschaft nachnutzbare Projektoutputs sind u.a. die Richtlinie zum Erstellen von KLARtexten, laienverständliche Einführungsvideos zum Thema “Metaanalyse”, sowie der KLARsaurus, ein Wörterbuch zu allgemeinverständlichen Übersetzungen von Fachbegriffen. Beteiligungsmöglichkeiten und community-basiertes Weiterentwicklungspotenzial werden aufgezeigt.

D-04: Lehrkräftebildung

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S06 S02 B06
Chair der Sitzung: Caroline Böning

Das Thema sprachliche Heterogenität in der Lehrkräftebildung aus der Perspektive der schulpraktischen Ausbildung (2. Phase)

Caroline Böning
Universität Duisburg-Essen, Institut für Deutsch als Zweit- und Fremdsprache

Die sprachliche Heterogenität von Lerngruppen zeigt sich auf diversen sprachlichen Ebenen in Form individueller, sprachlicher Repertoires (Busch 2013), die das Lernen von Fachinhalten und somit den schulischen Erfolg beeinflussen. Sprachliche Heterogenität war und ist nicht nur eine gesellschaftliche und schulische Realität an sich (Krüger-Potratz 2011) als Ausgangsbedingung für das Lernen, sondern auch Medium des Unterrichts, der spezifische sprachliche Anforderungen stellt. Dass sprachliche Heterogenität von Lerngruppen in jedem Unterricht berücksichtigt werden muss, ist mittlerweile kaum mehr umstritten (Gogolin 2020). Im Kontext sprachlicher Bildung schließt sich jedoch die Frage an, welche „Strategien des Umgangs“ (Niedrig 2002) Bildungs- und Schulsysteme sowie Lehrkräfte verfolgen, d. h. welche Sprache(n) in welcher Form mit welchen Motiven in Schule und Unterricht gefördert werden soll(en). Davon ausgehend, dass Professionalität von Lehrkräften durch die „wissens- aber auch erfahrungsgestützte produktive Verarbeitung beruflichen Erlebens [entsteht]“ (Cramer 2011), ist das Thema sprachlich heterogene Lerngruppen ein fester curricularer Bestandteil in allen Phasen der Lehrkräftebildung in Nordrhein-Westfalen. Im Gegensatz zur ersten Phase, dem Lehramtsstudium, (Schroedler & Lengyel 2018) gibt es in Bezug auf die zweite, schulpraktische Phase (Vorbereitungsdienst) der Lehrkräftebildung bisher kaum empirische Erkenntnisse. Dabei scheint es ein großes Forschungsdesiderat (Gerlach 2018) zu sein, an Konzepten wie dem Spannungsverhältnis zwischen Theorie und Praxis in der Lehrkräftebildung (Kunter 2011, Combe und Kolbe 2004) auch inhaltlich, hier hinsichtlich sprachlich heterogener Lerngruppen, anzuknüpfen und es phasenübergreifend zu betrachten.

Im Hinblick auf das Fehlen grundlegender empirischen Informationen zum genannten Thema im Vorbereitungsdienst stellt sich zunächst die Frage nach der grundsätzlichen Rolle, die dieses hier spielt: Wie wird das Thema in der zweiten Phase der Lehrkräftebildung im Rahmen der curricularen Vorgaben konzeptionell aufgefasst und ausgestaltet? Der Vortrag präsentiert für die Beantwortung dieser Frage eine Teilstudie eines Mixed-Methods-Forschungsprojekts mit einer schrittweisen Bearbeitung (Top-Down) dieser Fragen auf mehreren Ebenen (Meta und Meso). Nach der Darstellung curriculuarer Rahmenbedingungen für Nordrhein-Westfalen werden schulpraktische Perspektiven von Expert*innen des Feldes Vorbereitungsdienst präsentiert. Die Daten umfassen 13 leitfadengestützte, qualitative Expert*inneninterviews (Bogner et al. 2014) mit Seminarleitungen aus unterschiedlichen ‚Zentren für schulpraktische Lehrerausbildung‘ in Nordrhein-Westfalen. Sie wurden im Frühjahr 2022 erhoben und mit der inhaltlich-strukturierenden Inhaltsanalyse (Kuckartz 2014) ausgewertet. Im Vortrag werden die Ergebnisse von drei deduktiv-induktiv erstellten Hauptkategorien fokussiert: (1),Organisatorische Implementierung‘, (2),Inhaltliche Ausgestaltung/Strategien des Umgangs‘ sowie (3),Konzeptuelle Auffassung und institutionelles Wissen‘. Zentrale Ergebnisse der Analyse zeigen, dass für (1) das Thema sprachliche Heterogenität im Vorbereitungsdienst durchgängig entlang der Themen Unterrichtsplanung und -durchführung mitgedacht und auch durch Reflexion konkreter

Unterrichtssituationen, die von den Lehramtsanwärter*innen mit ins Seminar gebracht werden, behandelt wird. Auch bei (2) ergeben die Analysen, dass das Thema häufig mit anderen didaktisch-methodischen Themen verbunden wird. Am meisten Aufmerksamkeit erhält das Konzept sprachlicher Bildung, ‚sprachsensible Unterrichten', in Bezug auf fach- und bildungssprachliche Register, wohingegen die Themem ‚DaZ' und migrationsbedingte Mehrsprachigkeit eher unterschwellig thematisiert werden. Hinsichtlich der Kategorie (3) wird sprachliche Heterogenität als Normalfall und Realität wahrgenommen, jedoch in unterschiedlicher Art und Weise und Kontexten. Zentrale Themen sind diesbezüglich einerseits gesellschaftliche Veränderungen wie Migration. Wesentlich präsenter sind jedoch die regionalen Unterschiede der Schulstandorte und deren Auswirkungen auf sprachlich heterogene Lerngruppen der Lehramtsanwärter*innen, sowie der Rückbezug auf spezifische Schulformen. Sprachliche Bildung wird darüber hinaus unabhängig des Status ‚Migrationshintergrund', vor allem im Hinblick auf den Registererwerb wahrgenommen. Hinsichtlich der Vermittlung des Themas an die Lehramtsanwärter*innen spielen aus Sicht der Expert*innen neben dem Aufgreifen des Themas anhand von praktischen Unterrichtssituationen und dessen Reflexion sowie bedarfsorientiert vermittelten Methoden vor allem die Sensibilisierung und die Arbeit mit Haltungen hinsichtlich des Themas eine wichtige Rolle.

Basic Needs und Arbeitsprozessstrukturen als Determinanten für die Motivation zur Mitarbeit in phasenübergreifenden digitalisierungsbezogenen Communities of Practice der Lehrkräftebildung unter den Bedingungen der Corona-Pandemie

Anna Raneck-Kuhlmann, Jan Niemann, Heike M. Buhl, Birgit Eickelmann, Kerstin Drossel
Universität Paderborn, Deutschland

Die Förderung digitalisierungsbezogener Kompetenzen von Lehrkräften wird vor dem Hintergrund gesellschaftlicher Transformationsprozesse immer relevanter (Eickelmann et al., 2019). Dazu besteht ein hoher Bedarf an digitalisierungsbezogenen Ressourcen in der Lehrkräftebildung, die sowohl theoretisch fundiert als auch praxisorientiert sind. Ideal für die Entwicklung derartiger Ressourcen, ist die Zusammenarbeit von Vertreter*innen aller drei lehrkräftebildender Phasen in Communities of Practice (CoPs; van Ackeren et al., 2019). Da die Zusammenarbeit in CoPs als sozialer Lerngemeinschaft größtenteils auf freiwilliger Mitarbeit beruht, spielen die Motivation der CoP-Mitglieder und Möglichkeiten der Motivationsförderung eine wichtige Rolle bei einer solchen Zusammenarbeit.

Psychologische Ansätze verweisen hier auf die Relevanz der Basic Needs (Deci et al., 2017) und der Arbeitsprozessstrukturen (Bakker et al., 2007) für die Motivation zur Mitarbeit. Die Erfüllung der Basic Needs (Autonomie, Kompetenzerleben und sozialer Eingebundenheit) der Mitarbeitenden sowie eindeutige Arbeitsprozessstrukturen (Zielklarheit, Aufgabenkenntnis oder Informationsflussgestaltung) hängen positiv mit der Motivation zur Mitarbeit in Arbeitsgruppen zusammen (ebd.). Bisher unerforscht sind diese Zusammenhänge jedoch im Kontext der phasenübergreifenden Zusammenarbeit in der Lehrkräftebildung in CoPs. Hinzu kommen die Auswirkungen der Corona-Pandemie, die als Unsicherheitsfaktor mit einer massiven Adaption von Arbeitsprozessen in solchen Zusammenarbeitsformaten einhergeht und die bisher nur wenig erforscht wurden.

Der Vortrag bearbeitet dieses Forschungsdesiderat, indem folgende Forschungsfrage beantwortet wird: Welcher Zusammenhang zeigt sich zwischen der Erfüllung der Basic Needs sowie der Gestaltung der Arbeitsprozessstrukturen und der Motivation der beteiligten Akteur*innen zur Mitarbeit in phasenübergreifenden CoPs der Lehrkräftebildung auch unter Berücksichtigung der Herausforderungen der Corona-Pandemie?

Dazu werden Daten des vorhabenbegleitenden Qualitätsmanagements aus dem in NRW durchgeführten und vom Bundesministerium für Bildung und Forschung (BMBF) geförderten COMeIN-Verbundvorhaben (Communities of Practice für eine innovative Lehrerbildung; Projektlaufzeit 2020-2023, van Ackeren et al., 2019) verwendet. Das Projekt wurde vollständig unter den Bedingungen der Corona-Pandemie durchgeführt. Das COMeIN-Verbundvorhaben, an dem alle zwölf lehrkräftebildenden Universitäten NRWs sowie auch Vertreter*innen aus der zweiten und dritten Phase der Lehrkräftebildung beteiligt sind, hat u. a. zum Ziel digitalisierungsbezogene Ressourcen für das Lehrkräfteaus- und Fortbildungssystem zu entwickeln.

Im Oktober 2021 wurde eine schriftliche Online-Befragung mit allen CoP-Beteiligten (Rücklaufquote: 42.6%; n=171) mit quantitativen und qualitativen Anteilen durchgeführt. Die Auswertung der Daten zur Beantwortung der Forschungsfrage erfolgte mittels einer multiplen Regressionsanalyse und qualitativer Inhaltsanalyse.

Im Ergebnis der Regressionsanalyse zeigt sich ein signifikanter Zusammenhang zwischen dem Kompetenzerleben (ß= .348), der sozialen Eingebundenheit (ß= .200) und den Arbeitsprozessstrukturen (ß= .310) mit der Motivation der Mitglieder. Die Autonomie wies keinen zusätzlichen signifikanten Zusammenhang mit der Motivation auf (ß= -.105). Es konnte eine Varianzaufklärung von .48 (korr. R^2) erreicht werden.

Die qualitativen Ergebnisse verdeutlichen darüber hinaus, dass ein motivationsförderlicher Zusammenhang u. a. mit der Wahrnehmung der Relevanz des Projektes für die Lehrkräftebildung, einem persönlichen Interesse am Thema und einer gelingenden phasenübergreifenden Zusammenarbeit besteht. Als motivationsschwächend wurden u. a. fehlende zeitliche Ressourcen oder die Unklarheit bzgl. der Ziele und Aufgaben benannt. Auch scheinen sich die Bedingungen durch die Corona-Pandemie teilweise schwächend auf die Motivation ausgewirkt zu haben. So bestand einerseits der Wunsch nach mehr Präsenztreffen, wenngleich auch Vorteile für die Betonung der Relevanz des Themas Digitalisierung durch die Corona-Pandemie wahrgenommen wurden. Anderseits gaben Umfrageteilnehmende die Situation der Corona-Pandemie als Grund für ihre niedrige Motivation an.

Die Ergebnisse haben insofern eine hohe Bedeutung, als dass sie Steuerungswissen erzeugen, von dem sowohl das COMeIN-Verbundvorhaben als auch ähnliche Projekte der Lehrkräftebildung profitieren können. Hierbei bedarf es jedoch noch weiterer Forschung, beispielsweise welche langfristigen Veränderungen sich durch die Corona-Pandemie in Bezug zur Motivation in phasenübergreifenden CoPs der Lehrkräftebildung ergeben und inwiefern sich die Einflüsse auf die Motivation im Verlauf des Projektes verändern.

Unsicherheit gemeinsam durch Service Learning in der Lehrkräftebildung begegnen? Gelingensbedingungen für den reflektierten Einsatz von Virtual Reality in Schule und Unterricht

Helena Sträter[1], Caterina Schäfer[2], David Wiesche[3]
[1]Bergische Universität Wuppertal, Deutschland; [2]Universität Duisburg-Essen, Deutschland; [3]Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund

Der Transformationsprozess des Bildungssystems in der digitalen Welt (KMK, 2021) stellt an aktive und angehende Lehrkräfte die Anforderung, digitalisierungsbezogenen Kompetenzen weiterzuentwickeln und digitale Lehr-Lernformate reflektiert zu integrieren. Doch (angehende) Lehrkräfte fühlen sich unsicher in der Entwicklung digitaler Lernformate im Unterricht (Rubach & Lazarides, 2019).

Lehrkräftebildende Hochschulen müssen Rahmenbedingungen schaffen, in denen Studierende digitalisierungsbezogene Kompetenzen entwickeln können (Schäfer et al., 2021). Die KMK (2021) proklamiert, dass dabei immersive Technologien wie Virtual Reality (VR) in Schule zu beachten und zu reflektieren sind. Dazu fehlt es an Gelegenheiten zum Erproben und Reflektieren der Einsatzmöglichkeiten von VR in der Schule, der vor allem im prozess- und erfahrungsorientierten Lernen Potentiale ermöglicht (Radianti et al. 2020), gleichzeitig aus medizinischer, pädagogischer und ethischer Perspektive fehlende Evidenz und Unsicherheit berichtet wird (Zender et al. 2022). In diesem Spannungsfeld bewegen sich Lehrkräfte und Lehramtsstudierende gleichermaßen: sowohl in der Planung, Durchführung und Analyse von Lehr-Lern-Arrangements mittels VR sowie bezogen auf Professionalisierungs- und Schulentwicklungsprozesse sollen angehende Lehrkräfte handlungsfähig werden.

Die kompetenzorientierte Lehr- und Lernform *Service Learning* (Schank et al., 2020) ermöglicht einen aktiven und gestaltenden Umgang in phasenübergreifender Zusammenarbeit, in dem die Akteur*innen des Schul- und Bildungssystems übergeordnete gesellschaftliche Herausforderungen adressieren. So gelingt es, im Spannungsfeld Bildung und Digitalisierung angehende und aktive Lehrkräfte in praxisnaher Arbeit zusammenzubringen. Unklar ist jedoch, ob und wenn ja inwieweit das *Service Learning* auch für die Implementierung von technischen Innovationen wie VR in Schule sowie die kritisch-konstruktive Reflexion von didaktischen Einsatzmöglichkeiten zielführend ist.

Fragestellungen

Um diesem Desiderat zu begegnen, werden mit in einem Lehr- und Forschungsprojekt folgende Forschungsfragen adressiert:

• Welche Gelingensbedingungen lassen sich zur Implementierung von VR in Schule aus Perspektive von Studierenden und Lehrkräften abbilden?

• Welchen Beitrag leistet Service Learning dabei?

• Welche didaktischen Einsatzmöglichkeiten sehen Lehrkräfte für VR im Fachunterricht?

Als Intervention und konkrete Maßnahme der Schulentwicklung setzen sich Masterstudierende mit dem Querschnittsthema *Bildung in der digitalen Welt* und explizit mit der Bildungstechnologie VR (Radianti et al., 2020) auseinander. Studierende entwickeln gemeinsam mit Lehrkräften der

Partnerschulen ein bedarfsorientiertes Bildungsangebot mittels VR, führen es durch und reflektieren es im Sinne des Service Learnings. Der Prozess wird im Anschluss an die Lehrveranstaltungen im Sommersemester 2022 und Wintersemester 2022-2023 evaluiert.

Methoden

Zur Beantwortung der Forschungsfragen werden zum einen drei Ratingkonferenzen (Keller et al., 2012) mit Lehrkräften und Studierenden durchgeführt, um die Ergebnisberichte der Projektgruppen zu reflektieren.

Zum anderen werden zehn Leitfadeninterviews mit den Lehrkräften hinzugezogen, die hinsichtlich der Gelingensbedingungen von *Service Learning* zur Implementierung von VR sowie der kritische-konstruktiven Reflexion der didaktischen Einsatzmöglichkeiten Aufschluss geben.

Die Themen Phasenübergreifende Zusammenarbeit, Potentiale und Begrenzungen des Service Learnings sowie der mögliche Einsatz von VR im Fachunterricht und der Schule leiten sowohl die Ratingkonferenzen als auch die Interviews. Die Auswertung der erhobenen Daten erfolgt anhand Qualitativer Inhaltsanalyse (Kuckartz, 2018) mittels deduktiver und induktiver Kategoriensysteme.

Ergebnisse und ihre Bedeutung

Es werden Erkenntnisse zu Gelingensbedingungen der Implementierung von VR in Schule aus Perspektive von Studierenden und Lehrkräften im Kontext des *Service Learnings* präsentiert. Vorläufige Analysen zeigen, dass die Eigenerfahrung von VR von zentraler Bedeutung ist und eine grundsätzliche Bereitschaft der Lehrkräfte zur Implementierung in Schule besteht, die gleichzeitig Unsicherheiten mit sich trägt: Wird auf die Potentiale der Technologie im Unterricht hingewiesen, werden meist gleichzeitig einschränkende Kommentare mitformuliert („Ja, aber...“). Die Akteur*innen schätzen das Format Service Learning als zielführend ein, wenngleich sie neben dem gemeinsamen Lernweg auch VR-Expert*innen fordern. Weiterhin werden Erkenntnisse von didaktischen Einsatzmöglichkeiten im Fachunterricht vorgestellt, z.B. zeichnen sich im Sport u.a. Potentiale zum Training koordinativer Fähigkeiten ab.

So zeigt dieser Beitrag evidenzbasierte Erkenntnisse für die Gestaltung einer digitalisierungsbezogenen Lehrkräftebildung auf und eröffnet Perspektiven für die Auflösung von noch bestehenden Unsicherheiten.

Lehramtsstudium – Vorbereitungsdienst – Lehrkräfteberuf, oder? Empirische Analysen aus dem Nationalen Bildungspanel (NEPS) zum post-universitären beruflichen Verbleib von Lehramtsabsolvent*innen

Seyma Gülen[1], Svenja Mareike Schmid-Kühn[2], Katharina Müller[1]
[1]Leibniz Universität Hannover; [2]Universität Koblenz

Wenngleich der Großteil der Lehramtabsolvent*innen nach dem Studium den Vorbereitungsdienst aufnimmt (Wolf, 2015; Wieschke et al., 2017), entscheiden sich auch einige gegen das institutionelle Verlaufsmuster und schlagen (Berufs-)wege außerhalb des Lehramtes ein. Aufgrund der bisherigen Datenlage ist allerdings die Befundlage zum beruflichen Verbleib von Lehramtsabsolvent*innen bislang unzureichend, insbesondere hinsichtlich alternativer Verbleibsmuster außerhalb des klassischen, institutionell vorgegeben Weges von Lehrkräften. Hier setzt der Forschungsbeitrag an

und untersucht den post-universitären Verbleib von Lehramtsabsolvent*innen in einem Zeitraum von 10 Jahren nach Studienbeginn. Als Datengrundlage dient das Nationale Bildungspanel (NEPS), welches in der Startkohorte 5 Studierende seit ihrem Studieneintritt im Wintersemesters 2010/11 im Längsschnitt befragt (Blossfeld & Roßbach). Somit erlauben die NEPS Daten einerseits den spezifischen Blick auf den Übergang nach dem Lehramtsstudium. Andererseits geraten auch diejenigen nicht aus dem Blick, die sich vom klassischen Verlaufsmuster abwenden und nach erfolgreichem Lehramtsstudium einen anderen Berufs- bzw. Lebensweg einschlagen. Im Beitrag wird erstens der Forschungsfrage nachgegangen, wie sich der post-universitäre berufliche Verbleib von Lehramtsabsolvent*innen beschreiben lässt. Zweitens zielen die Analysen darauf ab aufzuzeigen, inwiefern sich der berufliche Verbleib von Lehramtsabsolvent*innen hinsichtlich soziodemographischer, leistungsbezogener und ausbildungsbezogener Merkmale unterscheidet. Anhand Methoden der bivariaten Statistik werden verschiedene post-universitäre Verbleibsmuster identifiziert, zu denen der professionsspezifische, akademische und der professionsunspezifische Verbleib zählen. Diese werden hinsichtlich soziodemographischer, leistungsbezogener und ausbildungsbezogener Merkmale vergleichend analysiert und charakterisiert (t-Test/chi2-Test). Die Ergebnisse zeigen in einer Gesamtschau, dass die überwiegende Mehrheit der Lehramtsabsolvent*innen erwartungskonform in den Lehrkräfteberuf einsteigt. Allerdings zeigen die Befunde auch post-universitäre Verläufe, die sowohl für bestimmte Formen des Lehramtes als auch für bestimmte individuelle Merkmale dieses klassische Verbleibsmuster durchbrechen. So schlagen insbesondere Lehramtsabsolvent*innen der Sekundarstufe II eine akademische Laufbahn ein, wohingegen sich ein professionsunspezifischer Berufseinstieg eher bei Berufsschullehramtsabsolvet*innen beobachten lässt. Wenn zusätzlich Leistungsindikatoren, wie die Note der Hochschulzugangsberechtigung oder die Note des Studienabschlusses betrachtet werden, ist eine positive Selektion in den akademischen Berufsverbleib festzustellen. Darüber hinaus zeigt sich bei Lehramtsabsolvent*innen, die in die Wissenschaft gehen, überdurchschnittlich häufig eine lineare Schullaufbahn vor Beginn des Studiums, weshalb sich seltener Auf- bzw. Abstiege während ihrer schulischen Bildungslaufbahn identifizieren lassen. Anhand der ausgewerteten Daten erlauben die Befunde einen Einblick hinsichtlich des post-universitären Verbleibs von Lehramtsabsolvent*innen sowohl im Längsschnitt innerhalb eines Zeitrahmens von 10 Jahren nach Studienbeginn, als auch differenziert nach Lehramtstypen und individuellen Merkmalen. So gelingt es mit den Analysen spezifische Gruppenmerkmale, die ein bestimmtes Verlaufsmuster begünstigen, zu identifizieren. Darüber hinaus lassen sich Aussagen über das Ausmaß sowie die Charakteristika unterschiedlicher post-universitärer Pfade von Lehramtsabsolvent*innen darlegen.

D-05: Inklusive Schule: Gemeinsamer Unterricht

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S06 S03 B06
Chair der Sitzung: Martina Osterrieder

Die Manifestation von Inklusion als Bildungs- und Erziehungsziel in den Lehrplänen bayerischer allgemeinbildender Schulen. Eine qualitative Inhaltsanalyse.

Martina Osterrieder, Anne-Christine Banze
Otto-Friedrich-Universität Bamberg, Germany

Im Rahmen einer qualitativen Inhaltsanalyse der in den fächerübergreifenden Richtlinien in den bayerischen Lehrplänen formulierten Bildungs- und Erziehungszielen wird der Indikator „Umsetzung inklusiver *Bildungsziele*“ (vgl. Steinmetz et al. 2021) beleuchtet.

Leitfragen für die vorliegende Analyse sind: Welche Konzepte von Inklusion manifestieren sich in den Bildungs- und Erziehungszielen der unterschiedlichen Schularten? Wird Inklusion als Begriff in der jeweiligen Schulart im Sinne eines spezifischen oder allgemeinen Konzeptes (vgl. Werning 2014, 606 ff.) verwendet? Welche Differenzkriterien von „Heterogenität“ finden in den Lehrplantexten Erwähnung? Wie werden inklusionsbezogene Werte in den Lehrplantexten kontextualisiert?

Forschungsstand und theoretischer Hintergrund:

Exemplarisch wird für die Untersuchung mit Bayern ein Bundesland gewählt, das zu den strukturpersistenten Ländern gerechnet wird und fast fünfmal so viele Schüler:innen mit sonderpädagogischem Förderbedarf an separaten Einrichtungen unterrichtet wie die restlichen Industriestaaten (vgl. Steinmetz et al. 2021, S. 75; S. 212). Mit den Bildungs- und Erziehungszielen, die spezifisch für jede Schulart jedem Lehrplan vorangestellt sind, werden Texte mit einer „Leit- und Prägefunktion für das gesamte Bildungs- und Schulsystem“ (Mißling & Ückert 2014, S. 38) analysiert. Die Lehrplantexte werden dabei als Ausdruck gesellschaftlicher Erwartungen an Schule (vgl. Künzli 1986; Künzli et al. 2013; Biehl et al. 1999; Terhart 2021) verstanden. Zwar lässt diese Wahl der Datengrundlage keine Aussagen über die Praxis an Einzelschulen zu, wohl aber über das kollektive Verständnis von Bildungszielen, die das Handeln an Schulen leiten *sollen*.

Werte werden in diesem Zusammenhang als situationsübergreifende Zielorientierungen und Entscheidungskriterien für das Urteilen und Handeln von Menschen wie auch Gruppen oder Institutionen, wie in diesem Fall der Schule, bzw. dem Ministerium, das die Lehrpläne verantwortet, verstanden. Werte werden von Individuen, Gruppen und Institutionen priorisiert und hierarchisiert, wenn Werturteile gefällt werden. (vgl. Schwartz 1999, S. 24 f.). Für die vorliegende empirische Analyse der inklusionsbezogenen Bildungsziele und den damit verbundenen Werten wurde die *refined theory of basic individual values* (Schwartz et al. 2012) zugrunde gelegt.

Methode:

Zur Beantwortung dieser Fragestellungen wurde eine strukturierende qualitative Inhaltsanalyse (Kuckartz 2014; Kuckartz & Rädiker 2022) durchgeführt. Datengrundlage sind die fachübergreifenden Richtlinien zu den Bildungs- und Erziehungszielen in den Lehrplantexten der vier allgemeinbildenden Schularten Grundschule, Mittelschule, Realschule und Gymnasium in Bayern.

Die Texte wurden zunächst in MAXQDA (Kuckartz, 2014) importiert und vorab in Analyseeinheiten eingeteilt. Eine Analyseeinheit bildet in der Regel ein vollständiger Satz, eine Überschrift oder seltener ein Stichpunkt einer Auflistung. Insgesamt bestehen die Kapitel aus N = 876 Segmenten

(Grundschule: *N* = 358; Mittelschule/Hauptschule *N* = 178; Realschule *N* = 187; Gymnasium *N* = 153.) Insgesamt handelt es sich um ca. 50 Textseiten.

Die Kategorisierungen haben zwei trainierte Personen unabhängig voneinander auf Basis eines Kategorienhandbuches (Osterrieder & Banze 2022) vorgenommen. Die Intercoderübereinstimmung lag bei der Werteanalyse bei κ=.90 und bei der Inklusionsanalyse bei κ=.84. Nur diejenigen Einheiten, die von den unabhängigen Codierenden in gleiche Kategorien eingeordnet wurden, wurden weiterverarbeitet. Neben einer deskriptiven quantitativen Auswertung wurde auch eine Analyse der Kontexte der Belegstellen sowie der Bedeutungshöfe vorgenommen.

Ergebnisse und ihre Bedeutung:

Im Ergebnis zeigen sich deutliche Unterschiede zwischen den Schularten. Während in den fächerübergreifenden Richtlinien für das Gymnasium der Begriff „Inklusion“ gar nicht und für die Realschule kaum explizit erwähnt wird, wird für Grund- und Mittelschulen der Begriff nicht nur expliziert, sondern in unterschiedlichen Kontexten anhand von Diskurslinien wie der Wertschätzung von Vielfalt hinsichtlich unterschiedlicher Heterogenitätsdimensionen entfaltet. Die Analyse der Inklusionskonzepte zeigt, dass die ungleiche Verteilung von Bildungschancen in den Lehrplänen der verschiedenen Schultypen systematisch reproduziert wird: Statt auf Chancengerechtigkeit zu zielen, nehmen die Lehrpläne die soziokulturelle und sozioökonomische Zusammensetzung der Schülerschaft vorweg. Daher werden diese Befunde schließlich vor dem Hintergrund von Fragen der Bildungsgerechtigkeit sowie dem Erfordernis der Behindertenrechtskonvention, ein inklusives Bildungssystem „auf allen Ebenen“ zu gewährleisten (vgl. Steinmetz et al. 2021), diskutiert.

Schulische Inklusion 2014 und 2022: Wie verändern sich Bedingungen und Intentionen inklusiven Verhaltens bei Lehrkräften über die Zeit?

Henrike Mentel, Natalie Förster, Elmar Souvignier
WWU Münster, Germany

Theoretischer Hintergrund

Inklusion wird – nicht zuletzt seit Inkrafttreten der Behindertenrechtskonvention 2009 – in zunehmendem Maße in der schulischen Praxis realisiert. Als Hauptakteure beeinflussen Lehrkräfte und ihre individuellen Merkmale maßgeblich die Umsetzung von Inklusion im Klassenraum (z.B. Savolainen et al., 2012). Gleichzeitig gilt, dass Lehrkräfte bei der Umsetzung mit anhaltenden strukturellen, gesellschaftlichen und persönlichen Veränderungen konfrontiert sind. Neben den Lehrkräften hängt die Umsetzung von Inklusion ebenfalls von schulischen Rahmenbedingungen ab (z.B. Woodcock & Woodcock, 2019). Bisherige Studien zeigen beispielsweise einen positiven Einfluss von Schulklima oder kollegialer Kooperation auf Einstellungen (Urton et al., 2014; Avramidis & Norwich, 2002) und Selbstwirksamkeit (Urton et al., 2014; Wilson et al., 2020; Hosford & O’Sullivan, 2016). Ziel des vorliegenden Beitrags ist daher, zu untersuchen, inwiefern sich Intentionen zur Umsetzung inklusiven Verhaltens und deren Bedingungen im Laufe der Zeit verändert haben und welche Rolle schulische Rahmen- und Kooperationsbedingungen spielen.

Als theoretisches Rahmenmodell zur Untersuchung dieser Frage bietet sich die Theory of Planned Behavior (TPB; Ajzen, 1991) an. Die Theorie postuliert, dass Verhalten maßgeblich von einer

entsprechenden Verhaltensintention abhängt, die wiederum von Einstellungen, subjektiver Norm und wahrgenommener Verhaltenskontrolle (Ajzen, 1991) beeinflusst wird. Im inklusiven Kontext wird die Theorie zunehmend eingesetzt, um inklusives Verhalten bzw. inklusive Verhaltensintentionen vorherzusagen (MacFarlane & Woolfson, 2013; Emmers et al., 2020; Lübke et al. 2016; Kuyini & Desai, 2007; Stanovich & Jordan, 1998; Ahmed et al. 2014; Opoku et al. 2021). Bisherige Studien konnten jedoch keine einheitlichen Prädiktoren der inklusiven Verhaltensintention identifizieren und Modelle unterscheiden sich in der aufgeklärten Varianz. Als Gründe sind neben Unterschieden in den Stichproben, den Operationalisierungen der Konstrukte und den Studiendesigns ebenfalls die Zeitpunkte der Studien zu nennen. Dies könnte einen Hinweis darauf geben, dass je nach Erhebungszeitpunkt und Rahmenbedingungen die inklusive Verhaltensintention durch unterschiedliche Prädiktoren vorhergesagt wird.

Generell kann man also von folgendem Wirkungsgefüge ausgehen: (Schulische) Rahmenbedingungen verändern sich und wirken sich auf Einstellungen, subjektive Norm und wahrgenommene Verhaltenskontrolle aus. Diese Veränderungen beeinflussen wiederum die inklusive Verhaltensintention. In der vorliegenden Studie soll dieser Wirkprozess im Inklusionskontext untersucht werden, wobei insbesondere mögliche Veränderungen des Wirkgefüges über die Zeit im Fokus stehen.

Fragestellungen

1. Haben sich Einstellungen, subjektive Norm und wahrgenommene Verhaltenskontrolle der Lehrkräfte zwischen 2014 und 2022 im Durchschnitt verändert?
2. Hat sich der Einfluss der Einstellungen, subjektiven Norm und wahrgenommenen Verhaltenskontrolle auf die Verhaltensintention im inklusiven Kontext im Laufe der Zeit verändert?
3. Hat sich der Einfluss der Rahmen- und Kooperationsbedingungen auf die Einstellungen, subjektive Norm und wahrgenommene Verhaltenskontrolle im Laufe der Zeit verändert?

Methode und Ergebnisse

Lehrkräfte unterschiedlicher Schulformen wurden in einer Online-Umfrage zu ihren Einstellungen, subjektiver Norm, wahrgenommener Verhaltenskontrolle, Verhaltensintentionen und schulischen Rahmen- und Kooperationsbedingungen befragt. Es werden zwei unterschiedliche Stichproben untersucht: Die erste Stichprobe wurde 2014 erhoben (N=192; Munske, 2015). Hier zeigte sich, dass die Skalen hohe interne Konsistenzen aufweisen (α = .72 - .93). Das spezifizierte TPB-Modell blieb zwar unter den empfohlenen Richtwerten für einen guten Modellfit (gemäß Hu & Bentler, 1999), zeigte jedoch einen akzeptablen Fit (RMSEA = .057, CFI = .890). Die Einstellungen stellten sich als alleiniger, signifikanter Prädiktor für die Verhaltensintention heraus (β = .60), die wiederum von den räumlichen, sachlichen, finanziellen und personellen Ressourcen der Schule (β = .36) und dem Ausmaß persönlicher Kooperationserfahrung (β = .29) beeinflusst waren.

Die Datenerhebung für die zweite Stichprobe begann im August 2022 und soll im Oktober abgeschlossen werden. Um die Vergleichbarkeit sicherzustellen, wurden dieselben Instrumente eingesetzt. Nach Untersuchung der Messinvarianz soll die Veränderung des Einflusses mittels einer multi-group SEM Analyse ausgewertet werden. Im Rahmen der Umsetzung von Inklusion können die Ergebnisse wertvolle Ansatzpunkte für Interventionen liefern und Aufschluss darüber geben, welchen Einfluss zeitliche Veränderungen und Rahmenbedingungen auf Verhaltensintentionen und ihre Prädiktoren im Kontext von Inklusion haben.

Profile multiprofessioneller Kooperation an inklusiven Schulen in Brandenburg

Franziska Rogge[1], Karsten Krauskopf[2], Antje Ehlert[1], Anne Hartmann[1], Jenny Lenkeit[1], Nadine Spörer[1], Michel Knigge[3]
[1]Universität Potsdam, Deutschland; [2]Fachhochschule Potsdam; [3]Humboldt-Universität zu Berlin

Die Zusammenarbeit von Lehr- und Fachkräften stellt im Kontext inklusiver Schul- und Unterrichtsentwicklung ein bedeutsames Element dar (Mora-Ruano et al., 2018). In einem organisationspsychologischen Verständnis (Spieß, 2004) ist Kooperation gekennzeichnet durch den Bezug auf andere, auf gemeinsam zu erreichende Ziele und bedarf Kommunikation sowie Vertrauen untereinander. Auf diesem Verständnis aufbauend differenzieren Gräsel et al. (2006) verschiedene Formen der unterrichtsbezogenen Kooperation (Austausch, Arbeitsteilung, Kokonstruktion) in Abhängigkeit gegebener Rahmenbedingungen und Intensität der Zusammenarbeit aus.

Vorliegende Forschungsbefunde verweisen darauf, dass Kooperation im schulischen Kontext häufig in wenig komplexen und eher autonomieerhaltenden Formen praktiziert wird (Muckenthaler, 2021; Richter & Pant, 2016), wobei dies von einzelschulspezifischen (Holzberger & Schiepe-Tiska, 2021) wie individuellen Voraussetzungen (Rogge et al., 2021) abhängig ist. Jene Unterschiede ermöglichen jedoch keinen Einblick hinsichtlich potenzieller qualitativer Unterschiede in der Zusammenarbeit. Inwiefern sich Lehr- und Fachkräfte, und auch Schulen, also anhand typischer Kooperationsmuster gruppieren lassen, ist bislang wenig erforscht (Holzberger & Schiepe-Tiska, 2021). Dies wäre jedoch wichtig zu eruieren, um u.a. Prozesse der Schulentwicklung (Prenger et al., 2018) und der individuellen Professionalisierung (Graham, 2007) bestmöglich zu unterstützen. Hierfür bedarf es jedoch komplexer personen-zentrierter Forschungsdesigns, um neben Profilen von Lehr- und Fachkräften auch Schulprofile identifizieren zu können.

Das Ziel des vorliegenden Beitrags ist es daher in einem personen-zentrierten Ansatz (Woo et al., 2018) auf der Basis der verschiedenen unterrichtsbezogenen Kooperationsformen (Gräsel et al., 2006) qualitativ unterschiedliche Profile von Lehr- und Fachkräften sowie darauf aufbauend Profile von Schulen zu explorieren. Daraus ergeben sich folgende Fragestellungen: (a) Inwiefern können distinkte Profile von Lehr- und Fachkräften auf der Basis ihres individuellen Kooperationsverhaltens identifiziert werden? (b) Inwiefern lassen sich distinkte Kooperationsprofile von Schulen eruieren und wie sind die individuellen Profile von Lehr- und Fachkräften in ihnen verteilt?

Basierend auf den Daten der „Evaluation des Gemeinsamen Lernens und Schulzentren im Bundesland Brandenburg“ (Lenkeit et al., 2021) konnten 2423 Lehr- und Fachkräfte (83% weiblich) der 184 inklusiv arbeitenden Schulen im Land Brandenburg, die im Frühjahr 2019 online befragt wurden, in die Analysen einbezogen werden. Zur Erfassung der realisierten Zusammenarbeit wurden die Skalen von Gräsel et al. (2006) mit den Dimensionen Austausch ($\alpha = .85$), Arbeitsteilung ($\alpha = .79$) und Kokonstruktion ($\alpha = .78$) verwendet. Da bislang wenig über die Existenz spezifischer Profile von Lehr- und Fachkräften wie auch Schulen hinsichtlich ihres Kooperationsverhaltens bekannt ist, wurden latente Profilanalysen in einem Mehrebenendesign (MLPA) (Mäkikangas et al., 2018; Nylund-Gibson et al., 2014) berechnet. Anschließend wurden ANOVAS und Tukey-post-hoc-tests (Nanda et al., 2021) zur Prüfung signifikanter Unterschiede zwischen den Profilen durchgeführt.

Es konnten vier Profile von Lehr- und Fachkräften identifiziert werden, die sich jeweils signifikant hinsichtlich des Kooperationsverhaltens unterscheiden: „sporadisch-kooperierend“, „unregelmäßig-kooperierend“, „regelmäßig-kooperierend“ und „intensiv-kooperierend“. Das größte Profil machten die „regelmäßig-kooperierenden“ Lehr- und Fachkräfte aus. Analysen auf Schulebene eruierten drei Profile von Schulen, die als „low-collaboration-schools“, „middle-collaboration-schools“ und „high-collaboration-schools“ bezeichnet werden können. Jene Schulprofile spiegeln dabei auch die Verteilung der individuellen Profile in ihnen wider. So findet sich in hoch kooperierenden Schulen bspw. auch mit ca. 60% das individuelle Profil der intensiv-kooperierenden Lehr- und Fachkräfte wieder. Neben individuellen Fähigkeiten wird folglich auch die Bedeutsamkeit der institutionellen Ebene deutlich, um Prozesse multiprofessioneller Kooperation, z.B. durch Schaffung eines kooperativen Klimas, zu unterstützen. Im Beitrag werden die Profile von Lehr- und Fachkräften sowie die der Schulen detailliert beschrieben und vor dem Hintergrund inklusiver Schul- und Unterrichtsentwicklung diskutiert.

Bedingungen und Bedarfe hinsichtlich Gemeinsamen Unterrichts an Berufsschulen

Simone Rückert, Helena van Vorst
Universität Duisburg-Essen, Germany

Ausgangslage

Berufskollegs sind ein Ort besonders ausgeprägter Diversität der Schüler*innen, an dem sich Lernende mit den unterschiedlichsten Eingangsvoraussetzungen versammeln. Dies resultiert unter anderem aus den Zugangsvoraussetzungen für eine Berufsausbildung: jede Person, die einen Vertrag mit einem Betrieb vorweisen kann, wird zugelassen, unabhängig von der vorherigen Schulbildung (Ministerium für Schule und Weiterbildung, Wissenschaft und Forschung des Landes Nordrhein-Westfalen 2022). Aufgrund der weiter steigenden Zahlen an Schüler*innen mit Zuwanderungsgeschichte, dem demographischen Wandel und dem Fachkräftemangel, stehen Berufskollegs und deren pädagogisches Personal vor der Aufgabe, Gemeinsamen Unterricht für alle zu ermöglichen (Euler und Severing 2014; IT NRW 2020; Heinrichs und Reinke 2019). Doch welche konkreten Förderbedarfe bestehen überhaupt bei Schüler*innen am Berufskolleg und wie gehen Lehrkräfte mit der Heterogenität der Schülerschaft am Berufskolleg um? Diesen Fragen soll in diesem Forschungsprojekt nachgegangen werden.

Methode und Instrumente

Das Vorhaben wird von folgender Forschungsfrage geleitet: Welche Unterstützungsbedarfe lassen sich bei der Schülerschaft von Berufskollegs diagnostizieren und welche Rahmenbedingungen herrschen an Berufskollegs im Hinblick auf Gemeinsamen Unterricht? (Schüler*innen- und Lehrer*innenperspektive)

Dazu wurden zunächst N=162 Schüler*innen von zehn Berufskollegs in Nordrhein-Westfalen im Bereich der dualen Ausbildungen zum Friseur/zur Friseurin und zum Kosmetiker/zur Kosmetikerin befragt. Der paper-pencil basierte Fragebogen enthält Likert-skalierte Items auf einer vierstufigen Skala zur Einschätzung der Zugänglichkeit des Unterrichts. Erhobene Kontrollvariablen waren u.a. der Migrationshintergrund, der selbsteingeschätzte Förderbedarf, Berufswahlmotive oder die

selbsteingeschätzten Sprachfähigkeiten in Deutsch. Die dazugehörigen Lehrkräfte (N=17) wurden mithilfe einer Online-Umfrage, bestehend aus offenen und wenigen geschlossenen Fragen, ebenfalls in die Untersuchung eingebunden. Es werden unter anderem Fragen zu Rahmenbedingungen, Förderbedarfen von Schüler*innen oder Diagnosemöglichkeiten gestellt. Die Auswertung der Lehrkräfte-Umfrage erfolgt mithilfe der Grounded Theory Methodology (GTM) (Glaser und Strauss 1967) in der Software MAXQDA 2022. Ein Ziel der Studie ist die Zusammenführung beider Perspektiven.

Ergebnisse und Diskussion

Mithilfe einer Faktorenanalyse der Likert-skalierten Items zur Zugänglichkeit des Unterrichts konnten vier Skalen identifiziert werden: Sprache & Verständnis (α = .86), Motivation & Interesse (α = .83), Unterrichtsinhalte (α = .84) und Unterrichtsgestaltung (α = .77). Ein selbsteingeschätzter Förderbedarf, der sich signifikant negativ auf alle ermittelten Skalen auswirkt, ist der einer psychosomatischen Problematik der Schüler*innen (häufige Kopf- und Bauchschmerzen, Schlafstörungen, Müdigkeit, schlechte Laune). Eine ANOVA ergab folgende Werte: Sprache & Verständnis - $F(1,160) = 14.62$, $p = .000$, $\eta^2 = .083$; Motivation & Interesse - $F(1,157) = 12.36$, $p = .001$, $\eta^2 = .073$; Unterrichtsinhalte - $F(1,159) = 16.08$, $p = .000$, $\eta^2 = .092$; Unterrichtsgestaltung - $F(1,160) = 6.08$, $p = .015$, $\eta^2 = .037$). Eine sprachliche Barriere der Schüler*innen wirkt sich negativ auf die Skala Sprache & Verständnis aus ($F(1,160) = 12.62$, $p = .001$, $\eta^2 = .073$), während ein selbsteingeschätztes Aufmerksamkeits-Defizit-Syndrom negative Auswirkungen auf die Skalen Sprache & Verständnis ($F(1,160) = 17.59$, $p = .000$, $\eta^2 = .099$) sowie die Zugänglichkeit der Unterrichtsinhalte hat ($F(1,159) = 9.63$, $p = .002$, $\eta^2 = .057$). Insgesamt sind die selbsteingeschätzten Beeinträchtigungen der Schüler*innen besonders im Bereich der emotional-sozialen Entwicklung (psychosomatische Beschwerden 43,8%), im Bereich der Aneignung und dem Merken von neuen Inhalten (21,0%) sowie bei sprachlichen Schwierigkeiten (21,0%) anzusiedeln. Die durch die Lehrkräfte benannten Bedarfe der Schüler*innen befinden sich vor allem in den Bereichen Sprachförderung, im Förderschwerpunkt Lernen, beim Konfliktmanagement sowie der Notwendigkeit einer stärkeren Verzahnung von Theorie und Praxis und decken sich somit zum Teil mit den durch die Schüler*innen selbst benannten Beeinträchtigungen. Zusammenfassend lassen zahlreiche Förderbedarfen und Beeinträchtigungen von Schüler*innen feststellen, die von Lehrkräften auch wahrgenommen werden. Gleichzeitig beklagt das Lehrpersonal schlechte Bedingungen für die Umsetzung Gemeinsamen Unterrichts. Ein möglicher Ansatz für die Hochschullehre und Praxisphasen könnte das frühe Heranführen von Lehramtsanwärter*innen an heterogene Klassen sein, um so die Selbstwirksamkeit und im Verlauf auch die Einstellung und Bereitschaft zum Unterrichten von diversen Klassen zu erleichtern.

D-06: Diversität: Schüler*innenleistungen und Ambitionen

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S06 S04 B06
Chair der Sitzung: Thomas Zimmermann

Ambitionen und soziale Herkunft: Bildungsentscheidungen zwischen Statusaufstieg und Statuserhalt

Birgit Becker, Thomas Zimmermann
Johann Wolfgang Goethe-Universität, Germany

Theoretischer Hintergrund und Fragestellung

Zur Erklärung sozial differenzierter Bildungsentscheidungen an den Gelenkstellen des Bildungssystems sind soziologische Theorien rationaler Wahlen (RAT) verbreitet (Breen und Goldthorpe 1997; Erikson und Jonsson 1996; Esser 1999). Nach diesen bewerten Schülerinnen und Schüler (SuS) die zur Verfügung stehenden Bildungsalternativen und wählen die Alternative mit dem höchsten subjektiven Erwartungsnutzen aus. Ein zentrales Motiv in diesem Abwägungsprozess ist der familiäre Statuserhalt, d. h. SuS sind bestrebt, nur solche Alternativen auszuwählen, die intergenerationell einen Statusabstieg vermeiden. Demzufolge ist die Möglichkeit eines Statusaufstiegs weniger relevant (vgl. Stocké 2005, 2012). Die geringere Relevanz eines Statusaufstiegs wird durch die Annahmen der Prospect Theory erklärt, wonach potenzielle Verluste schwerer wiegen als prospektive Gewinne (Kahneman und Tversky 1979). Ob die für die soziologische RAT zentrale Annahme des Statuserhalts als motivationaler Erklärungskomponente sozial differenzierter Bildungsentscheidungen zutreffend ist oder ob im Entscheidungsprozess der SuS nicht auch der Motivation für einen Statusaufstieg eine besondere Bedeutung zukommt, ist weitestgehend ungeklärt. Unsere Fragestellung lautet daher:

Welche Bedeutung haben Statuserhalts- und Statusaufstiegsmotivation für die Bildungsentscheidungen Jugendlicher in Familien mit unterschiedlichem sozioökonomischem Status (SES)? Variiert der Zusammenhang zwischen dem Streben nach Statuserhalt bzw. -aufstieg und dem Übergang in die Sekundarstufe II nach der sozialen Herkunft?

Methode

Datengrundlage ist die Startkohorte 4 (Klassenstufe 9) des Nationalen Bildungspanels (NEPS-Netzwerk 2021). Für den Übergang von der Sekundarstufe I in die Sekundarstufe II stehen Daten von insgesamt N=3.879 SuS zur Verfügung. Das Statusstreben der Jugendlichen operationalisieren wir durch die Differenz zwischen ihrem Wunschberuf und dem höchsten Beruf ihrer Eltern (jeweils gemessen durch den Internationalen Sozioökonomischen Index des beruflichen Status (ISEI)). Dabei unterscheiden wir zwischen Jugendlichen, die einen Beruf mit ähnlichem ISEI anstreben wie der Beruf ihrer Eltern (Statuserhalt), die einen Beruf mit höherem ISEI anstreben (Statusaufstieg) und die einen Beruf mit niedrigerem ISEI anstreben (Statusabstieg). Zur Analyse des Übergangs in die Sekundarstufe II verwenden wir lineare Wahrscheinlichkeitsmodelle mit robusten Standardfehlern.

Ergebnisse und ihre Bedeutung

Unsere Analysen bestätigen soziale Ungleichheiten beim Übergang in die Sekundarstufe II. SuS, deren Eltern in die höchste HISEI-Kategorie fallen (HISEI 65-max.) weisen unter Berücksichtigung der aktuell besuchten Schulform und Noten sowie weiterer soziodemografischer Variablen signifikant eine höhere Übergangswahrscheinlichkeit in die Sekundarstufe II auf als SuS anderer Herkunftsgruppen.

Hingegen ist die Wahrscheinlichkeit für einen Übergang von SuS der niedrigsten SES-Gruppe am geringsten. Aber auch das Statusstreben der SuS erweist sich am Übergang als relevant. Im Vergleich zu SuS, die den beruflichen Status ihrer Eltern erhalten möchten, zeigen SuS, die einen Statusaufstieg anstreben, höhere Übergangswahrscheinlichkeiten in die Sekundarstufe II und SuS mit geringeren beruflichen Ambitionen (Statusabstiege) niedrigere Übergangswahrscheinlichkeiten.

Bei genauerer Betrachtung der Rolle des Statusstrebens für verschiedene SES-Gruppen zeigt sich folgendes Bild: Unter den ausschließlich nach Statuserhalt strebenden SuS zeigen sich die eingangs berichteten Unterschiede im Übergangsverhalten, wonach mit einer höheren sozialen Herkunft ansteigende Übergangswahrscheinlichkeiten einhergehen. Streben SuS der mittleren und niedrigen SES-Gruppe aber einen Statusaufstieg an, dann weisen sie eine ähnliche Übergangswahrscheinlichkeit auf wie SuS der höchsten SES-Gruppe, die den beruflichen Status ihrer Eltern erhalten möchten.

Dieser Befund ist bedeutsam, da hiermit eine zentrale Annahme der soziologischen RAT infrage gestellt wird (Statuserhalt ist wichtiger als Statusaufstieg). In den niedrigeren und mittleren SES-Gruppen dominieren Berufswünsche, die einen intergenerationellen Statusaufstieg erwarten lassen und dieses Streben nach Statusaufstieg erweist sich für diese Gruppen als wesentlich motivierender für die hier untersuchte Bildungsentscheidung als ein Streben nach Statuserhalt.

Führt Tracking zur Sozialen Segregation oder zu homogenen Lernumwelten? Eine Diff-in-Diff-Analyse basierend auf den Schulleistungsstudien TIMSS, PIRLS und PISA

Maximilian Brinkmann[1], Nora Huth-Stöckle[2], Reinhard Schunck[2], Janna Teltemann[1]
[1]University of Hildesheim, Germany; [2]Universität of Wuppertal, Germany

Hintergrund der Studie

Die Auswirkungen von leistungsbasierter Stratifizierung („Tracking") auf Schulleistungen und soziale Bildungsungleichheiten sind Gegenstand anhaltender Debatten, wobei empirische Befunde tendenziell auf negative Effekte für Bildungsgleichheit und Nulleffekte für das allgemeine Leistungsniveau hindeuten (Terrin & Triventi 2022; Hanushek & Wößmann, 2006; Matthewes, 2020; Strello et al., 2021; Van de Werfhorst & Mijs, 2010). Zugleich stehen manche ForscherInnen diesen Befunden skeptisch gegenüber (Betts, 2011; Esser, 2016). Diese Debatte ist Ausgangspunkt unserer Studie, in der wir die zentralen Annahmen zweier konkurrierender Argumentationsstränge empirisch überprüfen.

Auf der einen Seite der Debatte steht die zentrale Annahme, Tracking führe durch die leistungsbezogene Eingruppierung der Kinder zu homogenen Lerngruppen. Hierdurch könne Tracking insbesondere über zwei Mechanismen positiv auf die Leistungen wirken: Organisationseffekte und Peer-Effekte (Hanushek & Wößmann 2006; Esser 2016; Brunello & Checchi, 2007). Durch Organisationseffekte können Curriculum und Unterricht spezifisch auf die homogenen Lerngruppen abgestimmt werden. Peer-Effekte beschreiben die Annahme, dass Kinder in homogenen Leistungsgruppen besser lernen können.

H1: Tracking führt zu einer stärkeren Leistungshomogenität.

Die andere Seite argumentiert hingegen, dass Tracking nicht strikt meritokratisch erfolge, sondern durch den sozialen Hintergrund beeinflusst sei. Erstens erfolge die Fähigkeits-Beurteilung beim Schulübergang indirekt über die beobachtete Leistung und sei durch das Elternhaus überformt ("primäre Effekte", Boudon, 1973), da Kinder aus privilegierten Elternhäusern bspw. besser auf die Schulumwelt vorbereitet würden. Darüber hinaus sei auch das Entscheidungsverhalten der Eltern beim Übergang durch den sozialen Status geprägt ("sekundäre Effekte"; Boudon 1973; Erikson & Jonsson, 1996; Breen & Goldthorpe, 1997). Als Folge ist bei gleicher Leistung der Übergang auf eine hohe Schulform (bspw. Gymnasium) für Kinder aus privilegierten Elternhäusern wahrscheinlicher. Zweitens legen empirische Befunde nahe, dass bei gleicher Leistung der soziale Status der Eltern einen Einfluss auf die Evaluierung der Leistung bzw. Übergangsempfehlungen durch die Lehrkräfte hat (Stubbe et al., 2011; Dumont et al., 2014). Als Folge dieser Mechanismen führe Tracking nicht zu der gewünschten Leistungshomogenität an Schulen, sondern zur sozialen Segregation der Schulen.

H2: Tracking führt zu einer stärkeren sozialen Segregation.

Obwohl die Leistungshomogenisierung und soziale Segregation zentrale Mechanismen sind, die in der Tracking-Debatte angeführt werden, wurden diese Annahmen selten empirisch getestet. In der vorliegenden Studie adressieren wir diese Forschungslücke und untersuchen den Effekt von Tracking auf die Leistungshomogenität und auf die soziale Segregation.

Daten und Methoden

Um unsere theoretischen Annahmen zu testen, stützen wir uns auf die Schulleistungsstudien PIRLS, TIMSS (4. und 8. Klasse) und PISA. Unsere Analysen basieren auf 78 Ländern, von denen je nach Operationalisierung von Tracking 20 bis 29 zu den getrackten Ländern gehören.

Um die Auswirkungen des Trackings zu ermitteln, verwenden wir einen Difference-In-Differences (DiD)-Ansatz (Wing et al., 2018). Wir nutzen dabei die Tatsache, dass kein Land die Leistungsgruppierung bereits in der Grundschule vollzieht. Da wir an zwei abhängigen Variablen interessiert sind (d.h. Leistungshomogenität und soziale Segregation), berechnen wir Seemingly Unrelated Regressions (SUR), die die gleichzeitige Untersuchung zweier abhängiger Variablen ermöglichen (Zellner, 1963).

Um die Robustheit unserer Ergebnisse über alternative Modellspezifikationen sicherzustellen, führen wir zudem Multiverse-Analysen durch. Hierdurch können wir bewerten, ob unsere Ergebnisse von einzelnen datenanalytischen Entscheidungen abhängen oder robust gegenüber alternativen Entscheidungen sind (Steegen et al., 2016; Simonsohn et al., 2020).

Ergebnisse und nächste Schritte

In Einklang mit beiden Perspektiven deuten bisherige Analysen auf einen Tracking-Effekt auf Leistungshomogenität sowie soziale Segregation hin, wobei der Effekt auf die Leistungshomogenität stärker und robuster ist. Als nächsten Schritt untersuchen wir einen potentiellen Moderatoren: die Freiheit der Schulwahl durch die Eltern.

Why and how diversity in the classroom affects students' achievement: social cohesion as a mechanism

Chenru Hou[1,2], Camilla Rjosk[1], Georg Lorenz[1,3]
[1]Institut zur Qualitätsentwicklung im Bildungswesen (IQB); [2]Humboldt-Universität zu Berlin; [3]Universität Leipzig

Theoretical Background

The educational practice has become increasingly challenging with increasing diversity in Western societies, schools, and classrooms (Gabaldón-Estevan 2020). Empirical research shows that diversity in multiple dimensions can affect students' achievement, such as diversity in ethnic and socioeconomic backgrounds (e.g., Conway-Turner, Williams, and Winsler 2020; Min and Goff 2016). However, the direction of these effects is mixed, and the mechanism of these effects is unclear. Classroom cohesion, which can be defined as the perception of group integration and supportive behaviors for group members (e.g., Meer and Tolsma 2014; Braaten 1991), is assumed to be a major mechanism behind the effects of diversity (Brechwald and Prinstein, 2011). This assumption is based on the social identity and categorization theory: diversity tends to reduce cohesion by leading to conflicts (Blalock 1967) since the members of a diverse group cannot develop common identities due to differences in characteristics (Tajfel and Turner 2004). Notwithstanding, socialization theory and social learning theory claim that students with peers from diverse backgrounds can receive more possibilities to learn from other groups and construct a better classroom climate, which may benefit their achievement (Raabe and Wölfer 2019; Khmelkov and Hallinan 1999).

Research Questions

Overall, this study aims to find out whether classroom cohesion is a mechanism explaining the effects of diversity on students' achievement. Diversity refers to ethnicity, socioeconomic background, prior achievement, and gender. The precise research questions are:

1) Does diversity reduce classroom cohesion?

2) Does classroom cohesion positively affect students' achievement?

3) Does classroom cohesion mediate the effects of diversity on students' achievement?

Method

This study is based on the BiKS 8-14 dataset in grades 5 to 7 (Artelt et al. 2013), with 1492 students in 122 classrooms. The Simpson's Index (Simpson 1949) for categorical variables (ethnicity, gender) and the standard deviation for continuous variables (prior achievement, socioeconomic background) are used to measure diversity as an independent variable in grade 5. Two items for the perception of classroom integration (e.g., "Ich komme mit den anderen Kindern in der Klasse gut aus", $\alpha = 0.64$) and four items for supportive behaviors (e.g., "Klassenklima: In der Klasse halten wir alle zusammen.", $\alpha = 0.68$) measure classroom cohesion in grade 6. The dependent variable is students' language test score in grade 7. OLS regression models with clustered standard errors including control variables (e.g., the share of minority students) are used to examine the research questions in preliminary analyses. Multilevel structural equation modeling will be used by the conference time.

Results

Preliminary analyses examine the associations at the student level, which confirm that only gender diversity negatively affects perceived supportive behaviors while other diversity dimensions do not provide significant effects. Diversity was not related to the perception toward group integration

(research question 1). Cohesion regarding supportive behaviors provides positive effects on students' achievement one year later, but the perception towards group integration shows no significant effect (research question 2). Diversity of students' socioeconomic background and ethnicity are negatively associated with students' achievement, but both are insignificant after controlling for average prior achievement in classrooms. Also, the results from path analysis show that both perspectives of cohesion are not significant mediators between diversity and achievement at the individual level (research question 3).

Nevertheless, the overlap in different dimensions of diversity (Meyer et al. 2014; Lau and Murnighan 1998) and possible differences in the effects of school experiences for students from different socioeconomic backgrounds (Brand and Xie 2010) are not considered in these results yet. Future analyses will involve a multidimensional diversity index and study if diversity effects differ according to students' socioeconomic backgrounds.

Meta-analytic review: Effekte der Leistungsheterogenität in Schulklassen und Schulen auf Schüler:innenleistung

Camilla Rjosk[1], Georg Lorenz[1], Svenja Hascher[1], Chenru Hou[1], Oliver Lüdtke[2]
[1]Institut zur Qualitätsentwicklung im Bildungswesen (IQB); [2]IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Theorie

Die Frage nach Vor- und Nachteilen des Lernens in leistungsheterogenen Schulklassen wird seit geraumer Zeit diskutiert, rückte aber insbesondere vor dem Hintergrund steigender Heterogenität der Schüler:innenschaft (Gräsel et al., 2017) wieder verstärkt in das öffentliche und wissenschaftliche Interesse. Einerseits wird z. B. angenommen, dass die Anpassung von Unterrichtsinhalten und Lerntempi in heterogenen Gruppen erschwert ist, was sich negativ auf die Leistungsentwicklung auswirke (vgl. Chiu et al., 2017; Decristan et al., 2017). Andererseits werden Vorteile in heterogenen Gruppen betont, u.a. dass Schüler:innen voneinander als Rollenmodelle lernen und sich gegenseitig unterstützen können (Bandura, 1971; Walgenbach, 2017). Bisherige Studien kommen zu uneinheitlichen Befunden (z.B. Decristan et al., 2017; Kuzmina & Ivanova, 2018). Allerdings fehlt bislang eine Synthese des Forschungsstands. Die vorliegende Arbeit setzt sich deshalb zum Ziel, den Forschungsstand verschiedener Disziplinen zu Effekten der Leistungsheterogenität innerhalb von Schulklassen und Schulen auf Schüler:innenleistung mittels metaanalytischer Methoden systematisch aufzuarbeiten.

Methode

Die systematische Literatursuche fokussiert den Zeitraum von 1990 bis 2020. Berücksichtigt werden Veröffentlichungen in englischer oder deutscher Sprache, die Effekte der Leistungsheterogenität innerhalb von Schulen oder Schulklassen der Primar- und Sekundarstufe auf Testleistungen anhand von Daten groß angelegter Stichproben untersuchen. Die Suche in Datenbanken verschiedener Disziplinen (*ERIC, PsycInfo, Scopus, EconBiz*) führte zu 6.461 Treffern (Zeitschriftenaufsätze, Working Paper, Buchkapitel). Anhand spezifischer Selektionskriterien wurden die Titel und Abstracts gescreent, relevante Studien auf ihre Eignung analysiert und inhaltlich mittels Volltextanalysen kodiert. Dieser mehrstufige Auswahl- und Reviewprozess wurde anhand von Doppelratings

durchgeführt. Zum Zeitpunkt der Abstracteinreichung liegen erste deskriptive Analysen vor, bis zum Konferenzzeitpunkt sind metaanalytische Auswertungen geplant.

Vorläufige Ergebnisse

Nach den verschiedenen Screeningschritten verblieben N = 29 Studien, aus denen 202 Effektgrößen kodiert wurden. Die Erstautor:innen der Studien sind mehrheitlich den Disziplinen Erziehungswissenschaften (41%) und Ökonomie (31%) zuzuordnen. Die Studien verteilen sich etwa gleichmäßig auf die Primar- und Sekundarstufe und sind größtenteils längsschnittlich angelegt (79%). Die Effektgrößen entstammen mehrheitlich Mehrebenenanalysen oder Regressionsanalysen mit Standardfehlerkorrektur unter Verwendung von Kontrollvariablen (49%) sowie Fixed Effect Modellen (37%). Das am häufigsten verwendete Heterogenitätsmaß (92%) ist die Standardabweichung der Schüler:innenleistung innerhalb von Schulklassen (66%) oder Schulen bzw. Jahrgangsstufen (34%). Als abhängige Variable werden in den meisten Fällen Testleistungen in Mathematik (59%) sowie sprachlichen Fächern einschließlich Leseverständnis (32%) analysiert. Die deskriptive Betrachtung der Signifikanz von Effektgrößen zeigt, dass sich zwei Drittel der Effekte (63%) nicht gegen den Zufall absichern lassen. Statistisch bedeutsame Effektgrößen verweisen in der Hälfte der Fälle auf positive Heterogenitätseffekte und in der anderen Hälfte auf negative Effekte. Die genannten deskriptiven Befunde werden bis zum Konferenzzeitpunkt metaanalytisch ausgewertet. Weiterhin sind Moderatorenanalysen geplant, um die Abhängigkeit der Effektgrößen von der Anlage der Studien (u.a. Primar- versus Sekundarstufe, Klassen- versus Schulebene) und methodischen Aspekten (u.a. statistische Modellierung, Verwendung von Kontrollvariablen) zu beleuchten. Zudem werden methodische Aspekte der Vergleichbarkeit von Studienbefunden diskutiert, die aufgrund der Interdisziplinarität auf sehr unterschiedlichen Vorgehensweisen fußen. Schließlich wird der Beitrag eine Einordnung der metaanalytischen Befunde in die eingangs beschriebene Debatte zur Bedeutung der steigenden Heterogenität von Schüler:innen für Lernerfolge vornehmen.

D-07: Förderung und Unterstützung

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S05 T00 B59
Chair der Sitzung: Jessica Jaeuthe

Einflüsse des sonderpädagogischen Förderbedarfs und der Wahrnehmung der Lehrkräfte-Schüler-Beziehungen auf das Lernverhalten

Jessica Jaeuthe[1], Anne Hartmann[1], Jenny Lenkeit[1], Antje Ehlert[1], Michel Knigge[2], Nadine Spörer[1]
[1]Universität Potsdam; [2]Humboldt-Universität Berlin

Die Förderung des Lernverhaltens stellt eine wichtige Zielvorgabe in den Bildungsplänen dar (Schreyer-Mehlhop et al., 2011). Jedoch mangelt es an Studien, die die Entwicklung des Lernverhaltens und möglicher Einflussmerkmale längsschnittlich untersuchen. Theoretisch wird angenommen, dass eine positive Beziehung zu Lehrpersonen und Peers ein wichtiges Merkmal für die Entwicklung des Lernverhaltens darstellen könnte (Looser, 2019). Auch empirisch zeigte die Meta-Analyse von Lei, Cui und Chiu (2016), dass positive Beziehungen zwischen Lehrpersonen und Schüler:innen mit einem positiveren Verhalten der Schüler:innen zusammenhingen.

Vor dem Hintergrund von inklusiven Lerngruppen stellt sich die Frage, welche Bedeutsamkeit die Beziehung für die Entwicklung des Lernverhaltens von Schüler:innen mit Förderschwerpunkt emotional-soziale Entwicklung aufweist. So zeigten Van Bergen, Graham & Sweller (2020), dass vor allem Schüler:innen mit Verhaltensauffälligkeiten ihre Beziehung zu Lehrpersonen negativer wahr, als Schüler:innen ohne Verhaltensauffälligkeiten.

Insgesamt ergeben sich daraus die folgenden Forschungsfragen:

1) Wie entwickelt sich das Lernverhalten der Schüler:innen von der zweiten zur dritten Jahrgangsstufe?

2) Inwiefern ist die Entwicklung des Lernverhaltens von der wahrgenommenen Beziehung zur Lehrperson und zu Peers abhängig?

3) Inwiefern zeigen sich differentielle Einflüsse der wahrgenommenen Beziehung auf das Lernverhalten für Schüler:innen mit einem und ohne einen Förderschwerpunkt im Bereich emotional-soziale Entwicklung (FSP ESE)?

Die Datengrundlage bildeten Erhebungen im Rahmen einer Längsschnittstudie in inklusiven Schulen in Brandenburg (Lenkeit et al., 2021). Es wurden Daten von *N* = 1159 Schüler:innen zu zwei Messzeitpunkten (t1: Beginn der zweiten Jahrgangsstufe, t2: Mitte der dritten Jahrgangsstufe) ausgewertet.

Das Lernverhalten wurde zu beiden Messzeitpunkten mit 10 Items durch die jeweilige Klassenleitung eingeschätzt (z.B. *Der Schüler bzw. die Schülerin strengt sich an, um eine Aufgabe zu lösen*; 0 = *Verhalten tritt nie* auf bis 3 = *trifft häufig auf;* Cronbachs α = .95, Petermann & Petermann, 2006).

Für die Einschätzung der wahrgenommenen Beziehungen wurden zu t2 basierend auf Rauer und Schuck (2004) das Gefühl des Angenommenseins (z.B. *Meine Lehrer kümmern sich um mich*, Cronbachs α = . 78) und die soziale Integration (z.B. *Ich fühle mich in der Klasse wohl*, Cronbachs α = .79) mit jeweils 6 Items von den Schüler:innen auf einer 4-stufigen Likert-Skala (0 = *stimmt gar nicht* bis 3 = *stimmt genau*) eingeschätzt.

Der diagnostizierte oder vermutete FSP ESE wurde durch die Lehrperson angegeben.

Als Kontrollvariablen wurden das Geschlecht und der HISEI der Eltern berücksichtigt.

Erste Ergebnisse:

Erste Analysen zeigten, dass das Lernverhalten insgesamt hoch eingeschätzt wird und von t1 zu t2 signifikant abnimmt (M_{t1} = 2.23, SD_{t1} = 0.63, M_{t2} = 2.10, SD_{t2} = 0.71). Es zeigten sich positive Korrelationen zwischen dem Lernverhalten, dem Gefühl des Angenommenseins und der sozialen Integration. Schüler:innen mit einem FSP ESE fühlten sich signifikant weniger sozial integriert, unterschieden sich jedoch nicht hinsichtlich des Gefühl des Angenommenseins von Schüler:innen ohne diesen Förderbedarf.

Erste Regressionsanalysen zeigten folgendes:

1) Das Lernverhalten zu t2 konnte signifikant durch das Lernverhalten zu t1, das Geschlecht (zugunsten der Mädchen) und den sozio-ökonomischen Hintergrund vorhergesagt werden.

2) Wurden das Gefühl des Angenommenseins und die soziale Integration getrennt voneinander in das Modell einbezogen, zeigten sich jeweils signifikante Zusammenhänge mit dem Lernverhalten. Wurden beide Variablen gemeinsam im Modell berücksichtigt, fiel nur der Einfluss der sozialen Integration signifikant aus. Dieses Ergebnis spricht dafür, dass beide Merkmale bedeutsam für die Entwicklung des Lernverhaltens sind, durch eine hohe geteilte Varianz jedoch bei der Berücksichtigung beider Merkmale die Bedeutsamkeit der sozialen Integration stärker ausfällt.

3) Schließlich zeigten sich differenzielle Effekte (Interaktion zwischen FSP ESE und Gefühl des Angenommenseins). Für Schüler:innen mit einem FSP ESE fiel der Einfluss der Beziehung zur Lehrperson auf die Veränderung des Lernverhaltens schwächer aus im Vergleich zu Schüler:innen ohne FSP.

Effekte eines Interventionsprogramms zur Förderung der sozialen Partizipation auf die Einstellungen von Grundschulkindern gegenüber Peers mit sonderpädagogischem Förderbedarf

Marwin Felix Löper, Gamze Görel, Frank Hellmich
Universität Paderborn, Germany

Mit der UN-Behindertenrechtskonvention (2006) und dem Beschluss der Kultusministerkonferenz (2011) sind Grundschulen mit der anspruchsvollen Aufgabe konfrontiert, inklusiven Unterricht zu gestalten, an dem alle Kinder entsprechend ihren individuellen Lernvoraussetzungen in vollem Umfang sozial und fachlich teilhaben können. Als wichtige Voraussetzungen für eine gelingende soziale Partizipation gelten dabei Freundschaften, Interaktionen und positive Einstellungen zwischen allen Kindern (Koster, Nakken, Pijl & van Houten, 2009). Grundschulkinder mit einem sonderpädagogischen Förderbedarf (SPF) stellen allerdings eine Risikogruppe dar, die von einer niedrigen sozialen Partizipation in der inklusiven Grundschule betroffen ist: So verfügen Kinder mit SPF meist nur über wenige Freundschaften, sie sind vermehrt von Interaktionsprozessen ausgeschlossen und werden von den anderen Kindern in der Klasse nur wenig akzeptiert (Avramidis, Avgeri & Strogilos, 2018). Die Einstellungen von Grundschulkindern gegenüber Mitschüler*innen mit SPF – verstanden als „psychologische Tendenzen, die von einer Person gegenüber einem Einstellungsgegenstand in Form von Zuneigung oder Ablehnung zum Ausdruck gebracht werden“

(Eagly & Chaiken, 1993, S. 1) – stellen in diesem Zusammenhang einen zentralen Ansatzpunkt dar, um der sozialen Ausgrenzung von Kindern mit SPF zu begegnen (Rademaker, de Boer, Kupers & Minnaert, 2020). Interventionen im Unterricht haben sich dabei als probates Mittel erwiesen, um die Einstellungen von Grundschulkindern gegenüber Peers mit SPF zu verbessern (z. B. Freer, 2021). Besonders effizient sind Interventionsansätze vor allem dann, wenn sie bereits in einer frühen Phase der kindlichen Entwicklung eingesetzt werden, langfristig angelegt sind, alle Kinder einer Lerngruppe miteinbeziehen, in den regulären Unterricht integriert sind und die Kinder sowohl kognitiv, affektiv als auch behavioral involvieren (z. B. Vuorinen, Erikivi & Uusitalo-Malmivaara, 2019).

Im Rahmen unserer Studie haben wir ein Interventionsprogramm zur Förderung der sozialen Partizipation von Kindern im dritten und vierten Schuljahr entwickelt und dahingehend empirisch untersucht, ob und inwiefern die Einstellungen der beteiligten Kinder durch die eingesetzten Interventionsmaßnahmen positiv beeinflusst werden können. Hierzu nahmen n = 259 Kinder über einen Zeitraum von sechs Wochen (zwei Unterrichtsstunden à 45 Minuten pro Woche) in ihrem Klassenverbund an Aktivitäten teil, in denen ihnen auf spielerische Art und Weise soziale Kompetenzen und Möglichkeiten, wie sie freundschaftlich in der Klassengemeinschaft miteinander umgehen können, vermittelt wurden. Die Kinder der Kontrollgruppe (n = 244) erhielten kein solches Training. Vor und nach der Intervention wurden die Kinder auf der Grundlage von Fallvignetten gebeten, Fragebogenitems zu ihren Einstellungen gegenüber Mitschüler*innen mit SPF in der emotional-sozialen Entwicklung (4 Items; Beispielitem: „Ich würde mich freuen, mit Julia in der Pause zu spielen.“; $\alpha_{Prä/Post}$ = .88/.90) sowie im Bereich Lernen zu beantworten (4 Items; Beispielitem: „Ich würde mich freuen, wenn Markus mein Freund wäre.“; $\alpha_{Prä/Post}$ = .83/.85). Das Antwortformat auf den einzelnen Einstellungsskalen reichte jeweils von 1 („stimmt gar nicht“) bis 5 („stimmt genau“).

Durch eine Varianzanalyse mit Messwiederholung konnte ein signifikanter Effekt durch die Intervention zur Förderung der sozialen Partizipation auf die Einstellungen der Grundschulkinder gegenüber Peers mit SPF in der emotional-sozialen Entwicklung ($F(1,499) = 16,79$; $p \leq .001$) sowie im Bereich Lernen ($F(1,499) = 10,95$; $p \leq .001$) nachgewiesen werden. Demnach sind die Kinder der Experimentalgruppe nach der Intervention signifikant positiver gegenüber Mitschüler*innen mit SPF im Bereich Lernen ($M_{Prä}$ = 3,14; M_{Post} = 3,24; $p \leq .001$) und in der emotional-sozialen Entwicklung ($M_{Prä}$ = 2,75; M_{Post} = 3,07; $p \leq .001$) eingestellt als vor der Intervention. Bei den Kindern der Kontrollgruppe blieben die Einstellungen hingegen konstant oder nahmen sogar ab. Insgesamt verdeutlichen die Ergebnisse unserer Studie, dass die Einstellungen von Grundschulkindern gegenüber Peers mit SPF positiv durch ein Interventionsprogramm zur Förderung der sozialen Partizipation verbessert werden können und geben erste Hinweise darauf, wie die soziale Teilhabe von Kindern mit SPF im inklusiven Unterricht der Grundschule in Zukunft verbessert werden kann.

Fachkraft-Kind-Interaktion – Dimensionen und Zusammenhänge mit Merkmalen pädagogischer Fachkräfte und Krippengruppen

Franka Baron[1], Anja Linberg[1], Simone Lehrl[2], Dorothea Dornheim[3]
[1]Deutsches Jugendinstitut; [2]Pädagogische Hochschule Weingarten; [3]Otto-Friedrich-Universität Bamberg

Frühkindliche Bildung, Betreuung und Erziehung (FBBE) ist für Kinder unter drei Jahren (U3-Bereich) zu einer Lernumwelt geworden, die aufgrund von massivem Ausbau und höherer Inanspruchnahme

mehr Aufmerksamkeit gewonnen hat (Autorengruppe Bildungsberichterstattung, 2020). Studien belegen, dass qualitativ hochwertige FBBE und Fachkraft-Kind-Interaktionen als Teil davon essenziell für eine gute Entwicklung von Kindern sind (u.a. Burchinal et al., 2015; Melhuish et al., 2015). Auch Studien, die Gruppen mit Kindern unter drei Jahren in den Fokus nehmen, zeigen diese positiven Assoziationen (Cassidy et al., 2017; Ruprecht et al., 2016; Salminen et al., 2021).

Dabei stellt sich die Frage, welche Facetten Interaktionsqualität beinhaltet und von welchen Merkmalen diese abhängen. Forschungsresultate stimmen in Bezug auf Kinder über drei Jahren überein, dass drei Dimensionen von Fachkraft-Kind-Interaktionen unterschieden werden können (u.a. Pakarinen et al., 2010), in Bezug auf den U3-Bereich liegen hierzu unterschiedliche Resultate vor. So wiesen Studien darauf hin, dass zwei Dimensionen unterschieden werden können: Interaktionen, die Emotionen und Verhalten der Kinder begleiten und Interaktionen, die Kinder bei Lerngelegenheiten unterstützen (Bichay-Awadalla & Bulotsky-Shearer, 2021; Guedes et al., 2020; van Schaik et al., 2018). Slot et al. (2017) zeigten hingegen, dass – wie bezogen auf ältere Kinder – drei Dimensionen unterschieden werden können und unterteilen zusätzlich Interaktionen der Emotionsbegleitung und der Verhaltensunterstützung.

Des Weiteren verdeutlichen Studien, dass Fachkraft-Kind-Interaktionen durch Strukturmerkmale sowohl auf Ebene der Fachkräfte (z.B. Ausbildung), als auch auf Ebene der Gruppe (z.B. Gruppengröße) beeinflusst werden können (Barros et al., 2016; Pessanha et al., 2007; Pianta et al., 2016). Darüber, welche Facetten von Fachkraft-Kind-Interaktionsqualität in Deutschland im U3-Bereich unterschieden werden können und wie diese mit Strukturmerkmalen assoziiert sind, liegen bisher jedoch wenige Erkenntnisse vor.

Dieser Beitrag geht daher folgenden Fragen nach:

(1) Welche Dimensionen von Fachkraft-Kind-Interaktionen lassen sich im U3-Bereich unterscheiden?

(2) Welche Zusammenhänge zwischen Fachkraft-Kind-Interaktionen und Strukturmerkmalen zeigen sich im U3-Bereich auf Fachkraft- und Gruppenebene?

Die Fachkraft-Kind-Interaktionen wurde in 95 Krippengruppen mit CLASS Toddler (La Paro et al., 2012) an einem regulären Vormittag im Gruppenalltag beobachtet. Die Interaktionsqualität wurde in Hinblick auf acht Bereiche (Positives Klima, negatives Klima, Feinfühligkeit, Berücksichtigung der kindlichen Perspektive, Verhaltenslenkung, Unterstützung von Lernen und Entwicklung, Feedbackqualität, Sprachgebrauch) auf siebenstufigen Skalen eingeschätzt. Je eine Fachkraft aus der Gruppe füllte einen Fragebogen aus, durch den Strukturmerkmale erhoben wurden. Für die Analysen wurden Angaben zur Aus- und Weiterbildung, zur Berufserfahrung und zum Alter der Fachkräfte, zur Gruppengröße, zum Fachkraft-Kind-Schlüssel, zur Alterszusammensetzung der Kinder sowie zum Anteil von Kindern mit Migrationshintergrund in der Gruppe verwendet.

Um die Forschungsfragen zu beantworten, wurden konfirmatorische Faktorenanalysen (CFA) und Strukturgleichungsmodelle (SEM) in R berechnet. Resultate der CFA bestätigten, dass sich zwei Dimensionen von Fachkraft-Kind-Interaktionen im U3-Bereich unterscheiden lassen: Emotionale und verhaltensbezogene Interaktionen und aktiv lernunterstützende Interaktionen. Das Modell zeigte eine akzeptable Modellpassung und war passender als ein 1- oder 3-Faktoren-Modell: $\chi^2(19)=56.51$,

p<.001, CFI=.91, TLI=.87, RMSEA=.14, (CI90:.10-.19), SRMR=.07. Ergebnisse des SEM zeigten Assoziationen von drei Strukturmerkmalen mit der Fachkraft-Kind-Interaktionsqualität: Die Ausbildung und das Alter der Fachkräfte sowie die Alterszusammensetzung der Kinder innerhalb der Gruppe. Dieses Modell zeigte eine angemessene Passung: $\chi^2(43)=81.65$, *p*<.001, CFI=.91, TLI=.88, RMSEA=.10 (CI90:.06-.13), SRMR=.06. Geringfügig weniger emotionale und verhaltensbezogene Unterstützung wurde beobachtet, wenn die Fachkräfte älter waren (*B*=-.02, *SE*=.01, *p*=.01), stärkere negative Assoziationen mit dieser Dimension zeigte eine große Altersmischung innerhalb der Gruppe (*B*=-.35, *SE*=.15, *p*=.02). Eine etwas engagiertere Lernunterstützung leisteten Fachkräfte mit höherem Bildungsabschluss (*B*=.12, *SE*=.07, *p*=.10). Altersgemischte Gruppen waren negativ mit der aktiven Lernunterstützung verbunden (*B*=-.33, *SE*=.16, *p*=.04). Weitere Strukturmerkmale waren nicht mit den Dimensionen von Fachkraft-Kind-Interaktionen im U3-Bereich assoziiert.

Die Ergebnisse werden in Hinblick auf die Modellpassungen und die Implikationen für Forschung und Praxis diskutiert. Insgesamt trägt dieser Beitrag zu einem besseren Verständnis über Zusammenhänge zwischen verschiedenen Dimensionen von Fachkraft-Kind-Interaktionen und Strukturmerkmalen im U3-Bereich bei.

Neue Wege der digitalen Unterstützung von Familien mit Kleinkindern – Teilnahmemuster eines digitalen Familienprogramms während der Corona-Pandemie

Luisa Prokupek, Sabine Blaurock
Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund

Im Frühjahr 2020 erreichte die Corona-Pandemie Deutschland und stellte Familien in der Bewältigung ihres bisherigen Alltages vor große Herausforderungen. Eltern mit kleinen Kindern waren im Vergleich zu anderen Eltern von diesen Herausforderungen besonders betroffen, u.a. da Kinder dieser Altersgruppe eine intensivere Betreuung benötigen (Hübner et al., 2020). Maßnahmen zur Eindämmung des Corona-Virus hatten zur Folge, dass in Präsenz stattfindende Familienunterstützungsangeboten temporär wegfielen. Gerade Familien von sehr jungen Kindern verloren dadurch Möglichkeiten der Familienbildung sowie der Vernetzung. Darüber hinaus sind sehr junge Kinder in ihrer (sprachlichen) Entwicklung in hohem Maße von der anregenden häuslichen Lernumgebung (HLE) ihrer Familie (z.B. durch die Häufigkeit sprachförderlicher Aktivitäten, wie dem Vorlesen) abhängig (Linberg, 2017). Obwohl Studien positive Zusammenhänge von digitalen Elterninterventionen und der Stärkung der familialen Anregungsqualität belegen (Jelley, Sylva & Karemaker, 2016), nutzen bisher nur wenige Familienbildungsprogramme das Potenzial digitaler Medien (Ereky-Stevens et al., 2019). Während der Corona-Pandemie führten einerseits Kontaktbeschränkungen zu Familienmitgliedern, wie z.B. der Großelterngeneration, zu einer geringeren sozialen Unterstützung von Familien. Andererseits vermehrten sich Stressfaktoren durch Ängste, finanzielle Sorgen und der Vereinbarkeit von Kinderbetreuung und Beruf (Hübener et al., 2020). Familialer Stress und mangelnde soziale Unterstützung können sich wiederum negativ auf die Eltern-Kind-Interaktionen auswirken, wodurch die Anregungsqualität der HLE abgeschwächt werden kann (Masarik & Conger, 2017).

Fragestellung

Im Fokus des vorliegenden Beitrags stehen daher die Fragen, welche Eltern durch digitale Angebote erreicht werden können und welche Bedürfnisse teilnehmende Eltern während der Corona-Pandemie in Bezug auf die (sprachliche) Anregungsqualität, die empfundene soziale Unterstützung und das Stresserleben aufweisen.

Methode

Für die Beantwortung der Fragestellungen werden die Daten des Interventionsprojektes "DIGIFam - Entwicklung und Implementierung einer digitalen pädagogischen Komponente zur Steigerung der familialen Anregungsqualität" zum 1. Erhebungszeitpunkt vor der Intervention mit einem Stichprobenumfang von 221 Familien mit Kleinkindern (M = 13 Monaten) genutzt. Die Eltern wurden in zwei (nicht randomisierte) Interventionsgruppen (Gruppe A: Elternkurs und App, N = 81; Gruppe B: nur App, N = 79) sowie eine Kontrollgruppe (N = 61) unterteilt und im Hinblick auf ihre familialen Hintergrundmerkmale und ihrer Ausgangssituation bezüglich ihrer Bedürfnisse miteinander verglichen (Selektivitätsanalysen mittels multinominaler logistischer Regressionsanalysen).

Ergebnisse

Insgesamt weisen die Ergebnisse darauf hin, dass es sich um eine sozial selektive Stichprobe handelt und wenig bildungsbenachteiligte Personen teilnehmen. Für Familien mit nicht oder nicht ausschließlich deutscher Familiensprache und mittlerem Bildungsniveau war die Wahrscheinlichkeit höher der Gruppe A „Elternkurs und App“ als der Gruppe B „nur App“ anzugehören (OR = 8,6, p = .001). Differenzierte Teilnahmemuster zeigen sich bei der familialen Anregungsqualität und dem Stresserleben der Familien. Familien mit einem vergleichsweise zur Gruppe B „nur App“ hohen Entwicklungspotenzial in der familialen Anregungsqualität erhalten sowohl den Elternkurs als auch die App (Gruppe A) (OR = 0,8, p = .099). Im Vergleich hatten die Eltern der Gruppe B "Nur App" ein erhöhtes Stressempfinden: Je geringer das empfundene Stressniveau (OR = 0,6, p = .050), desto höher war demnach die Wahrscheinlichkeit, zu Gruppe A "Kurs und App" zu gehören. Dies weist auf das Potenzial digitaler Familienprogramme (wie Apps) für eine flexible, unabhängige Nutzung im herausfordernden Familienalltag hin.

Die Besonderheiten der HLE sind speziell in dieser sehr jungen Altersgruppe bisher wenig erforscht. Unsere Ergebnisse tragen dazu bei, mehr über die Bedürfnislage von Familien mit Kleinkindern während der Corona-Pandemie zu erfahren. Weiterführende Analysen über die insgesamt drei Erhebungszeitpunkte (prä-, post – und follow-up Online-Befragung) werden zeigen, ob die Intervention (und wenn ja, Kurs und/oder App) die Bedürfnisse der Familien zur Ausgangssituation aufgreifen und kompensatorische Effekte erzielen konnte. Die Ergebnisse werden im Hinblick auf das Potenzial digitaler Familienbildungsprogramme (Zeit- und Ortsunabhängigkeit) für Familien mit Kleinkindern diskutiert.

D-08: Belastung von Lehrkräften

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S05 T00 B71
Chair der Sitzung: André Meyer

Studentische Lehrkräfte in Schulen: eine Untersuchung ihrer schulpraktischen Tätigkeiten, ihrer sozialen Unterstützung und ihres Belastungsempfindens

André Meyer, Eric Richter, Sebastian Kempert
Universität Potsdam, Deutschland

Theoretischer Hintergrund

Angeleitete Praxisphasen im Lehramtsstudium bilden eine wichtige Komponente zur Entwicklung professioneller Kompetenzen angehender Lehrkräfte (Brouwer & Korthagen, 2005; Dreer, 2022). Um Praxiserfahrungen auszuweiten sowie aufgrund des hohen Bedarfs an Schulen gehen viele Studierende nebenberuflichen Tätigkeiten nach, in deren Rahmen sie einzelne Schüler*innen betreuen, als Vertretungslehrkräfte eingesetzt werden oder eigenständig Fachunterricht durchführen (vgl. Ulrich & Gröschner, 2020). Die beruflichen Erfahrungen und Tätigkeiten als studentische Lehrkräfte wurden bislang jedoch nicht systematisch untersucht.

Empirische Untersuchungen weisen auf eine hohe berufliche Belastung bei Berufseinsteiger*innen im Lehrberuf hin (Klusmann et al., 2012; Roloff et al., 2022). Entsprechend des Job-Demands-Resources Modells ist anzunehmen, dass insbesondere studentische Lehrkräfte u.a. soziale Unterstützung durch Kolleg*innen benötigen, um mit beruflichen Belastungen angemessen umgehen zu können (Jolly et al., 2020; Xanthopoulou et al., 2007). Aufgrund des dramatischen Lehrkräftemangels ist jedoch davon auszugehen, dass ihnen diese Ressource kaum zur Verfügung steht.

Während Evidenz zu den Effekten institutionalisierter Praxisphasen im Lehramtsstudium existiert (z.B. Praxissemester; Ulrich & Gröschner, 2020), liegen keine Befunde zur Gestaltung nicht-institutionalisierter schulpraktischer Tätigkeiten studentischer Lehrkräfte vor. Ferner ist bislang unbekannt, wie belastet sich studentische Lehrkräfte wahrnehmen und welche soziale Unterstützung sie durch Kolleg*innen erfahren. Der Beitrag verfolgt daher folgende Forschungsfragen:

1. Welchen unterrichtlichen Tätigkeiten gehen studentische Lehrkräfte in Schulen nach und wie hängen diese Tätigkeiten mit ihren demografischen und professionsbezogenen Merkmalen zusammen?
2. Wie nehmen studentische Lehrkräfte die soziale Unterstützung und ihr Belastungsempfinden in den Schulen wahr?
3. Besteht ein Zusammenhang zwischen den unterrichtlichen Tätigkeiten studentischer Lehrkräfte, ihrer sozialen Unterstützung und ihrem Belastungsempfinden?

Methode

Wir fragten eine Stichprobe von N = 107 Lehramtsstudierenden, die einer nebenberuflichen Tätigkeit an einer Schule nachgehen (85% weiblich, 51% Bachelorstudierende, 60% an Grundschulen tätig), welche Tätigkeiten sie an ihren Schulen übernehmen (z.B. Ich betreue einzelne Schüler*innen, Ich führe eigenständig Unterricht durch), wie sie die soziale Unterstützung durch Kolleg*innen wahrnehmen (z.B. Meine Kolleg*innen bieten mir Hilfe an, wenn ich diese benötige; ω = .89) und wie sie ihre berufliche Belastung einschätzen (z.B. Ich fühle mich wegen meiner beruflichen Belastung oft müde und angespannt; ω = .89).

Ergebnisse und ihre Bedeutung

Die Ergebnisse einer latenten Klassenanalyse deuten darauf hin, dass sich die befragten studentischen Lehrkräfte anhand ihrer Tätigkeiten in drei Gruppen differenzieren lassen. Bei Gruppe 1

(*n* = 37) handelt es sich um Personen, die vorrangig eher wenig komplexen Tätigkeiten nachgehen (z.B. einzelne Schüler*innen betreuen). Gruppe 2 (*n* = 43) setzt sich aus Personen zusammen, die mehreren unterschiedlich komplexen Tätigkeiten in der Schule nachgehen (z.B. Teamteaching, Vertretungsstunden übernehmen, eigenständig Unterricht durchführen). Gruppe 3 (*n* = 27) umfasst Personen, die hauptsächlich eigenständig Fachunterricht durchführen. Die Ergebnisse zweier Chi-Quadrat-Tests weisen auf signifikante Unterschiede zwischen den Gruppen hinsichtlich des Studienfortschrittes der Studierenden, $\chi^2(2) = 17.80$, $p < .001$, $\varphi = .41$, sowie der Schulform an der sie tätig sind hin, $\chi^2(2) = 12.92$, $p < .01$, $\varphi = .35$. Dabei fühlen sich die befragten studentischen Lehrkräfte in hohem Maße sozial unterstützt ($3.20 < M < 3.38$) und nehmen ein moderates Maß an beruflicher Belastung wahr ($2.13 < M < 2.67$). Die Ergebnisse einer multivariaten ANOVA weisen auf signifikante Unterschiede zwischen den Gruppen hinsichtlich des berichteten Belastungsempfindens hin ($F(2/93) = 3.75$ $p < .05$, $\eta^2 = .08$). Studentische Lehrkräfte, die eher weniger komplexen Tätigkeiten nachgehen berichten über eine geringere Belastung als studentische Lehrkräfte, die hauptsächlich selbstständig Unterricht durchführen. Mehrgruppenvergleiche in einem latenten Strukturgleichungsmodell zeigen für alle drei Gruppen einen negativen Zusammenhang zwischen der sozialen Unterstützung und dem Belastungsempfinden der studentischen Lehrkräfte. Somit scheint die soziale Unterstützung in der Schule unabhängig vom Grad der Komplexität der schulpraktischen Tätigkeiten eine wichtige Rolle für eine geringe wahrgenommene Belastung zu spielen.

Welche Rolle spielen Ressourcenveränderungen im Praxissemester für die erlebte Beanspruchung von angehenden Lehrkräften?

Katharina Neuber
Universität Duisburg-Essen, Deutschland

Der Lehrberuf gilt aufgrund der Komplexität und Vielfalt beruflicher Anforderungen als in besonderem Maße herausfordernd (Rothland, 2013). Dabei kann ein effektiver Umgang mit den eigenen Ressourcen die Bewältigung beruflicher Belange unterstützen und zu beruflichem Wohlbefinden beitragen (Buchwald & Hobfoll, 2004). Bereits vor Eintritt in die Schulpraxis werden angehende Lehrkräfte mit vielfältigen Anforderungen konfrontiert, die potenziell herausfordernd sein können und deren Bewältigung entsprechende Ressourcen erfordert. So gilt bspw. das Praxissemester im Lehramtsstudium als eine Phase erhöhter Beanspruchung, da angehende Lehrkräfte in Tätigkeitsfeldern aktiv werden, die sich von der universitären Ausbildung deutlich unterscheiden und zudem mit divergierenden Rollenerwartungen einhergehen (Schüle et al., 2017). Trotz der hohen Relevanz der beruflichen Beanspruchung für eine erfolgreiche Ausübung der Lehrtätigkeit (z. B. Klusmann et al., 2006) und empirischen Hinweisen darauf, dass bereits die im Studium erlebte Beanspruchung zum beruflichen Wohlbefinden in der späteren Schulpraxis beitragen kann (z. B. Zimmermann et al., 2016), steht die empirische Betrachtung der Beanspruchungssituation von Studierenden im Praxissemester bislang noch am Anfang (Vogelsang, 2021). Um den bisherigen Erkenntnisstand zu erweitern, werden im vorliegenden Beitrag längsschnittliche Befunde zur Relevanz wahrgenommener Ressourcenveränderungen für die Veränderung der erlebten Beanspruchung von Praxissemesterstudierenden präsentiert. Dabei stehen folgende Fragen im Fokus:

1. *Wie verändert sich das Beanspruchungserleben im Praxissemester und welche Ressourcenveränderungen nehmen Studierende wahr?*
2. *Welche Zusammenhänge zeigen sich zwischen den wahrgenommenen Ressourcenveränderungen und der Beanspruchung am Ende des Praxissemesters?*

Das Beanspruchungserleben von *N* = 124 Lehramtsstudierenden wurde zu Beginn und zum Ende des Praxissemesters mithilfe schriftlicher Fragebögen erfasst. Als Indikatoren der Beanspruchung wurden in Anlehnung an die Studierenden-Version des Maslach-Burnout-Inventory von Gumz et al. (2013) zum einen die *emotionale Erschöpfung* (6 Items, Bsp.: Ich fühle mich von meinem Praxissemester ausgelaugt; α t1/t=.90/.93); zum anderen die erlebte *Effizienz* der Studierenden erhoben (4 Items, Bsp.: Ich kann die Probleme, die in meinem Praxissemester auftreten, effektiv lösen; α t1/t2=.66/.69). Zum Ende des Praxissemesters sollten die Studierenden zudem einschätzen, inwieweit sie in der Praxisphase unterschiedliche Ressourcen gewonnen (+4=maximaler Ressourcengewinn) oder verloren (-4=maximaler Ressourcenverlust) haben. In Anlehnung an Marwinski und Buchwald (2010) wurden dabei die Bereiche *persönliche Ressourcen* (5 Items, Bsp.: Berufliche Motivation, α=.75), *Bedingungsressourcen* (6 Items, Bsp.: Unterstützung durch die Schulleitung, α=.81) und *Energieressourcen* (5 Items, Bsp.: Freizeit, α=.65) adressiert.

Im Rahmen von Mittelwertvergleichen zeigen sich erwartungskonform signifikante Veränderungen sowohl für die *emotionale Erschöpfung* (M_{t1}=2.27, M_{t2}=2.48; t(123)=-2.662, p=.004, d=0.239) als auch für das *Effizienzerleben* der befragten Studierenden (M_{t1}=4.27, M_{t2}=4.59; t(123)=-5.034, p<.001, d=0.452). Die deskriptive Auswertung der Ressourcenveränderungen verdeutlicht einen insgesamt eher geringen Ressourcengewinn im Praxissemester: Zwar überwiegen die positiven Ausprägungen der einzelnen Indikatoren, jedoch nehmen nur drei Items einen Wert über dem theoretischen Skalenmittelwert einer positiven Ressourcenveränderung (M=2,5) an. Dies trifft zu auf die *persönliche Ressource* „berufliche Motivation“ (M=2,56; SD=1,63; n=124) sowie auf die *Bedingungsressourcen* „Verhältnis zu den Schüler*innen“ (M=2,98; SD=1,09; n=124) und „positives Klimas im Kollegium“ (M=2,56; SD=1,62; n=124). *Energieressourcen*, insbesondere „Freizeit“ (M=-1.37; SD=1,64; n=122), wurden von den Studierenden hingegen verstärkt als bedroht wahrgenommen. In multiplen Regressionsanalysen konnte – unter Kontrolle des jeweiligen Ausgangswerts der Beanspruchung – die Relevanz der Ressourcenveränderungen für die erlebte Beanspruchung bekräftigt werden. So wiesen Studierende mit einem Zugewinn an *Bedingungsressourcen* am Ende des Praxissemesters eine geringere *emotionale Erschöpfung* (β=-0,187, p=.029) und ein höheres *Effizienzerleben* (β=0,328, p<.001) auf.

Der Mehrwert der Studie liegt darin, die protektive Wirkung unterschiedlicher Ressourcenbereiche für die Beanspruchung auch für den Kontext des Praxissemesters aufzuzeigen. Die Befunde bekräftigen insbesondere den Stellenwert, den soziale Ressourcen wie die Unterstützung durch Praxislehrpersonen für Studierende in Praxisphasen einnehmen, und legen zusammenfassend nahe, dass angehende Lehrkräfte hinsichtlich des Aufbaus persönlicher wie sozialer Ressourcen möglichst frühzeitig unterstützt werden sollten, um die vielfältigen Anforderungen der Lehrpersonenausbildung und späteren Schulpraxis erfolgreich zu bewältigen.

Andreas Ortenburger, Stefanie Gäckle, Claudia Menge
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), Germany

Dem hohen Bedarf an Lehrpersonal an allgemeinbildenden und beruflichen Schulen in Deutschland bei gleichzeitig niedriger Verfügbarkeit vollausgebildeter Lehrkräfte begegnet die Bildungspolitik aktuell zunehmend mit einer Öffnung der beruflichen Zugangswege zum Lehrkräfteberuf (Dedering 2020; Lucksnat et al., 2022). In vielen Bundesländern ist somit der Eintritt im sogenannten Quer- oder Seiteneinstieg möglich, d. h. ohne abgeschlossenes Lehramtsstudium und/oder ohne Lehramtsreferendariat. Dieser Öffnung steht die Unsicherheit gegenüber, ob trotz des Fehlens wichtiger Etappen des berufsbiographischen Entwicklungsprozesses (Terhart, 2011) dennoch hinreichend professionelle Kompetenzen zur Verfügung stehen, um sich im anspruchsvollen Lehrkräfteberuf bewähren zu können. Der hier antizipierte Mangel an berufsbezogenem Wissen nährt die Befürchtung einer niedrigeren Unterrichtsqualität von Quer- und Seiteneinsteiger*innen, auch wenn es dafür noch keine empirische Evidenz gibt (Ziegler et al., 2022). Bisherige Studien deuten teilweise auf vorhandene Unterschiede zwischen traditionell und nicht-traditionell ausgebildeten Lehrkräften z. B. hinsichtlich fachdidaktischen und erziehungswissenschaftlichen Wissens hin (z.B. Kleickmann & Anders, 2011; Lucksnat et al., 2020, 2022). Dies ist insofern plausibel, als das diese Kompetenzdimensionen zum einen eine gezielt wissenschaftsbasierte Auseinandersetzung mit Theorien des Lehrens und Lernens voraussetzen, und zum anderen eine reflektierte Anwendung in der Praxis, um professionsbezogenes Handlungswissen überhaupt erst generieren zu können (Oser, 2001; Terhart, 2003). Die Befundlage ist gleichwohl alles andere als eindeutig und zudem sind in jüngeren Untersuchungen oftmals keine Seiteneinsteiger*innen berücksichtigt oder zwischen Quer- und Seiteneinsteiger*innen kann nicht hinreichend unterschieden werden.

Hier setzt der geplante Beitrag an, der den Blick auf den Vergleich von Quer- und Seiteneinsteiger*innen richtet und Unterschiede hinsichtlich professionsbezogener Werthaltungen und Überzeugungen als weitere Kompetenzfacetten im generischen Modell professioneller Handlungskompetenz (Baumert & Kunter, 2006) zu Beginn ihrer beruflichen Laufbahn als Lehrkräfte untersucht. Als zentrale Dimensionen von Unterrichtsqualität gelten u. a. Klassenführung, kognitive Aktivierung und konstruktive Unterstützung (vgl. z. B. Klieme et al., 2006), die in diesem Beitrag im Vergleich zwischen Lehrkräften mit unterschiedlichen, nicht-traditionellen Zugangswegen zum Lehrkräfteberuf untersucht werden. Zudem wird ein Blick auf transmissive und konstruktivistische Überzeugungen zum Lehren und Lernen geworfen, die eine wichtige Rolle für das professionelle Handeln von (angehenden) Lehrkräften spielen (Kunter & Pohlmann, 2015).

Auf Basis der Daten der NEPS-Studierendenkohorte (NEPS-Netzwerk, 2022), in der zunehmend auch Personen identifiziert werden können, die als Quer- und Seiteneinsteiger*innen in Schulen unterrichten (n=132), wird untersucht, inwieweit sich Lehrkräfte mit unterschiedlichen, nicht-traditionellen Berufszugangswegen hinsichtlich professionsbezogener Überzeugungen und selbsteingeschätzter Handlungskompetenz unterscheiden, wobei davon ausgegangen wird, dass insbesondere Seiteneinsteiger*innen niedrigere Ausprägungen professioneller Kompetenzen beobachten lassen. Die abhängigen Dimensionen werden mit guten Messeigenschaften latent modelliert (am Beispiel professionsbezogener Überzeugungen mit zwei Subdimensionen:

Chi²(25)=33,98, P=0,108; CFI=0,966; TLI=0,951; RMSEA=0,951). Die Ergebnisse deuten lediglich in Bezug auf individualisierten Unterricht auf Unterschiede zwischen den beiden nicht-traditionell ausgebildeten Lehrkräftegruppen hin, wobei Seiteneinsteiger*innen zu Beginn ihrer beruflichen Karriere im Lehramt erwartungsgemäß weniger differenziert auf die individuellen Bedürfnisse ihrer Schüler*innen im Unterricht eingehen als Quereinsteiger*innen. Im Beitrag werden Anlage und Ergebnisse der Untersuchung dargestellt, Limitationen diskutiert und zukünftige Untersuchungsschritte abgeleitet.

Neue Zugänge zum Belastungserleben von Lehrkräften: Bieten Herzratenmessungen mit Fitnesstrackern eine effiziente, günstige und robuste Messmethode?

Mandy Klatt, Peer Keßler, Gregor Kachel, Christin Lotz, Anne Deiglmayr
Universität Leipzig

Theoretischer Hintergrund:

Physiologischen Daten wie z. B. die Herzfrequenz erhalten in der Forschung zum Stresserleben zunehmende Bedeutung. Sie stellen einen wichtigen Indikator für körperlichen oder emotionalen Stress dar, da eine erhöhte Arbeitsbelastung mit einer erhöhten Herzfrequenz einhergeht (Sachs, 2014). Weiterhin ermöglichen sie eine objektivere Erfassung von Stress als Selbstberichte (Runge et al., 2020). Jedoch setzt die Erfassung der Herzrate im Bildungskontext den Einsatz von kostengünstigen und non-invasiven Instrumenten voraus. Fitnesstracker, die am Handgelenk getragen werden, haben das Potential, ein solches nützliches Instrument zu sein (Ferguson et al., 2015).

Bisher gibt es jedoch nur wenig Evidenz über die Brauchbarkeit von Herzratenmessungen mittels Fitnesstrackern in Lehr- und Lernsettings (Ertzberger & Martin, 2016; Lowe, 2016). Alleinig Runge et al. (2021) untersuchten an einer verhältnismäßig kleinen Stichprobe (N = 4 Lehrpersonen) den Stress von Lehrpersonen und zeigten, dass eine hohe Herzfrequenz, hohe Schrittzahl und wenig Schlaf auf mehr Stress bei Lehrpersonen hinweisen.

Es besteht daher weiterhin ein Mangel an belastbaren Studien zu der Frage, ob sich Fitnesstracker als effiziente, günstige und robuste Messmethode zur Erfassung des Belastungserlebens von Lehrkräften eignen. Da unter anderem eine höhere Berufserfahrung mit geringerem Stressempfinden im Lehrerberuf einhergeht (Fisher, 2011), sollen in dieser Studie auch die Herzratenmessungen von erfahrenen und unerfahrenen Lehrpersonen betrachtet werden.

Fragestellungen:

Vor diesem Hintergrund untersuchten wir, ob Herzfrequenzmessungen mittels Fitnesstracker in einem kontrollierten Lehr-Lern-Setting geeignet sind, um Unterschiede zwischen

(1) Antizipations-, Erregungs- und Ruhephasen sowie

(2) zwischen erfahrenen und unerfahrenen Lehrpersonen abzubilden.

Dabei erwarteten wir, dass die Herzraten in der Erregungsphase höher ist als in den Ruhephasen und untersuchen explorativ, ob erfahrene Lehrpersonen im Vergleich zu unerfahrenen eine niedrigere Herzfrequenz aufweisen.

Methode:

Die geplante Stichprobe besteht aus *n* = 40 erfahrenen Lehrpersonen (Lehramtsstudium abgeschlossen, im Lehrberuf tätig) und *n* = 40 unerfahrenen Lehrpersonen (Lehramtsstudierende, erstes Praktikum absolviert). Die derzeit laufende Datenerhebung ist Teil eines größeren Forschungsprojektes zur professionellen Wahrnehmung von Unterrichtsstörungen. Bislang wurden Daten von *n* = 19 erfahrenen und *n* = 31 unerfahrenen Lehrpersonen erhoben. Die Datenerhebung wird bis zum Zeitpunkt der Konferenz abgeschlossen sein.

In einem kontrolliertem Laborsetting unterrichteten die Lehrpersonen eine 15-minütige, selbst vorbereitete Unterrichtslektion vor einer "Klasse" von drei Schauspielenden. Die gemittelten Herzfrequenzdaten in Schläge pro Minute (bpm) wurden mittels Fitbit Charge 4 über einen Gesamtzeitraum von zwei Stunden erfasst. Daraus wählten wir drei theoretisch interessierende 10-minütige Intervalle für die Analysen aus:

(1) Antizipationsphase vor Lektionsstart,

(2) Erregungsphase ab Lektionsstart und

(3) Ruhephase zum Abschluss der gesamten Erhebung.

Zur Beantwortung der Fragestellungen berechneten wir eine 3 (Phasen: Antizipation vs. Erregung vs. Ruhe) × 2 (Lehrerfahrung: erfahren vs. unerfahren) ANOVA (via lme4 in R).

Ergebnisse:

Vorläufige Analysen des bereits erhobenen Samples ergaben einen signifikanten Haupteffekt für die Phasen. Post-hoc Tests zeigten, dass die durchschnittliche Herzrate in der Antizipationsphase und in der Erregungsphase bedeutsam höher war als in der Ruhephase. Jedoch unterschied sich die Antizipationsphase nicht von der Erregungsphase.

Der Haupteffekt „Lehrerfahrung“ ergab keinen statistisch bedeutsamen Mittelwertsunterschied der durchschnittlichen Herzfrequenz zwischen den erfahrenen und unerfahrenen Lehrpersonen. Ebenso fand sich kein statistisch bedeutsamer Interaktionseffekt (Phasen × Lehrerfahrung), erfahrene und unerfahrene Lehrpersonen zeigten also einen vergleichbaren Verlauf über die drei Phasen hinweg.

Diskussion:

Anhand der Analysen wird deutlich, dass am Handgelenk getragene Fitnesstracker eine geeignete Methode sind, um im Laborsetting unterschiedliche Phasen von Erregungszuständen bei (angehenden) Lehrpersonen zu messen. Zukünftige Forschung könnte Lehrpersonen in ihrer alltäglichen Schulpraxis begleiten, um mögliche Unterschiede in der Berufserfahrung sichtbar zu machen.

Mittels der finalen Ergebnisse können Potentiale und Gelingensbedingungen von physiologischen Daten für die Unterrichtsforschung diskutiert und ein Beitrag für den Bildungskontext geleistet werden. So könnte der Einsatz von tragbaren Fitnesstrackern neue Einblicke in das Belastungserleben von Lehrpersonen und Lehramtsstudierenden liefern, um mögliche stressreduzierende Maßnahmen in der Lehramtsbildung zu implementieren.

D-09: Unterrichtsqualität II

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S05 T00 B83
Chair der Sitzung: Isabell Runge

Wie wirkt die Unterrichtsqualität mit subjektiven Wertüberzeugungen und der Leistung Studierender zusammen?

Isabell Runge[1], Rebecca Lazarides[1], Charlott Rubach[2], Luise von Keyserlingk[3], Jutta Heckhausen[4], Sandra Simpkins[4], Richard Arum[4], Jacquelynne S. Eccles[4]
[1]Universität Potsdam, Deutschland; [2]Universität Rostock, Deutschland; [3]Technische Universität Dortmund, Deutschland; [4]University of California, Irvine, United States

Theoretischer Hintergrund

Die Leistungen Lernender und ihr leistungsbezogenes Handeln werden unter anderem durch das Lehrkräftehandeln und subjektive Wertüberzeugungen der Lernenden beeinflusst – dies wird theoretisch von der Situated Expectancy-Value Theory (SEVT; Eccles & Wigfield, 2020) beschrieben und in Befunden verschiedener empirischer Arbeiten aufgezeigt (Bergey et al., 2018; Fauth et al., 2014; Holzberger et al., 2019; Nilsen et al., 2016; Roeser et al., 2000). So sind beispielsweise höhere Leistungen zu verzeichnen, wenn Lernende Inhalte als interessant und als für ihre Zukunft nützlich und wichtig empfinden. Zudem konnten Studien zeigen, dass das wahrgenommene Lehrkräftehandeln mit subjektiven Wertüberzeugungen von Lernenden in Zusammenhang stehen (Fauth et al., 2014; Ruiz-Alfonso et al., 2021; Schreier et al., 2014). Die Befunde zur Erklärung von Lernerleistungen und leistungsbezogenem Handeln beschränken sich zumeist auf den schulischen Kontext. Inwiefern diese Zusammenhänge jedoch im universitären Kontext bestehen, wurde bisher kaum untersucht.

Ziel und Fragestellung der vorliegenden Studie

Ziel der vorliegenden Studie ist es, die längsschnittlichen Zusammenhänge zwischen der wahrgenommenen Unterrichtsqualität, subjektiven Wertüberzeugungen sowie der Leistung von Studierenden im Verlauf eines Semesters zu untersuchen. Dabei werden die benannten Zusammenhänge im schwierigsten und wichtigsten Kurs der Studierenden analysiert, um die Kontextspezifizität der Wertüberzeugungen und ihrer Korrelate zu untersuchen (Eccles & Wigfield, 2020). Die Studie geht der Frage nach, inwieweit die von den Studierenden wahrgenommene Unterrichtsqualität mit Veränderungen der subjektiven Wertüberzeugungen und der Abschlussnote der Studierenden in ihrem schwierigsten und wichtigsten Kurs zusammenhängt.

Methode

Es wurden Daten von 428 Studierenden einer staatlichen Universität in den USA, Kalifornien, ausgewertet, die im Rahmen eines Semesterquartals mehrmals an einer Onlinebefragung teilnahmen. In der vorliegenden Studie wurden validierte Skalen zur Erfassung der Unterrichtsqualitätsdimensionen Klassenführung, kognitive Aktivierung und konstruktive Unterstützung (Messzeitpunkt 1; Rubach et al., 2022) sowie zur Erfassung der subjektiven Wertüberzeugungskomponenten Wichtigkeit, Interesse und Nützlichkeit (Messzeitpunkte 1 und 2; Eccles & Wigfield, 2002) genutzt und spezifisch für den schwierigsten und den wichtigsten Kurs erfasst. Die Unterrichtsqualität sowie subjektiven Wertüberzeugungen wurden jeweils als Second-Order-Factor operationalisiert. Zur Einschätzung der Leistung gaben die Studierenden ihre Abschlussnoten in ihrem schwierigsten und wichtigsten Kurs zum dritten Messzeitpunkt an. Zur

Auswertung der Daten wurden „Latent Change Score (LCS)"-Modelle mittels Mplus 8.7 modelliert (Muthén & Muthén, 1998-2021). Im ersten Schritt wurde die Messinvarianz der Skala zu den subjektiven Wertüberzeugungen zwischen den Messzeitpunkten getestet.

Ergebnisse

Die Ergebnisse der LCS-Modelle für die Forschungsfrage zeigten, dass die von den Studierenden wahrgenommene Unterrichtsqualität in ihrem *schwierigsten* Kurs in einem signifikant positiven Zusammenhang mit dem Intercept (β = .37, SE = 0.10, $p < .001$) sowie mit den Veränderungen der subjektiven Wertüberzeugungen (β = .24, SE = 0.09, $p = .034$) sowie mit der Abschlussnote (β = .24, SE = 0.09, $p = .031$) standen. Für den *wichtigsten* Kurs wiederum stand die wahrgenommene Unterrichtsqualität zwar mit dem Intercept der Wertüberzeugungen in positivem Zusammenhang (β = .47, SE = 0.10, $p < .001$), jedoch nicht mit den Veränderungen der subjektiven Wertüberzeugungen (β = -.09, SE = 0.07, $p = 0.356$) noch mit der Abschlussnote (β = .02, SE = 0.14, $p = 0.905$). Die Veränderungen der subjektiven Wertüberzeugungen im *wichtigsten* Kurs standen in einem signifikant positiven Zusammenhang mit der Abschlussnote (β = .23, SE = 0.09 $p = .020$), jedoch bestanden keine signifikanten Zusammenhänge im *schwierigsten* Kurs (β = .07, SE = 0.09 $p = .445$).

Relevanz

Die hier vorgestellten längsschnittlichen Daten geben Hinweis darauf, dass die von Studierenden wahrgenommene Unterrichtsqualität insbesondere in solchen Kursen für die subjektiven Wertüberzeugungen und die Leistung von Bedeutung ist, welche von Studierenden als besonders schwierig eingeschätzt werden. Damit zeigt sich, wie in der SEVT (Eccles & Wigfield, 2002) beschrieben, die Relevanz des Kontextes, in dem Studierende sich befinden.

Wie lief die Unterrichtsstunde? Entwicklung, Validierung und Erprobung eines Kurzfragebogens zur Unterrichtsqualität aus Sicht von Schüler:innen und Nutzung der Ergebnisse durch Lehramtsstudierende

Jan Pfetsch[1], Christine Ladehoff[1], Diemut Ophardt[1], Christiane Buchholtz[1], Carolin Lohse[1], Felicitas Thiel[2], Holger Gärtner[2]

[1]Technische Universität Berlin; [2]Freie Universität Berlin

Lehramtsstudierende entwickeln Unterrichtskompetenzen in praxisvorbereitenden Lerngelegenheiten (Futter, 2017; Ophardt, 2021), etwa beim eigenständigen Unterrichten im Praxissemester. In Unterrichtsnachbesprechungen werden Stärken, Schwächen und Handlungsalternativen der Unterrichtsdurchführung zumeist aus Sicht von Lehramtsstudierenden und Universitätsdozierenden reflektiert. Eine zusätzliche Perspektive auf die Unterrichtsqualität könnte jedoch das Feedback durch Schüler:innen zur jeweiligen Unterrichtsstunde bieten.

Datengestützte Entscheidungen über das Lehren und Lernen finden in der Lehrkräftebildung noch wenig Verbreitung. Dennoch zeigte sich, dass Feedback das Verständnis und Durchführung von professionellem Verhalten verbessern kann (Hattie & Timperley, 2007; Kluger & DeNisi, 1996) und dass die datenbasierte Reflexion die Wirkung von Unterricht auf den Lernprozess von Schüler:innen fördern kann (Hattie, 2008). Wenn sich die Wahrnehmung der Unterrichtsqualität zwischen Lehrkräften und Schüler:innen unterscheidet, das Feedback persönlich relevante Dimensionen des Unterrichts betrifft, und zumindest auch positive Rückmeldungen beinhaltet, dann sind Lehrkräfte

besonders geneigt, ihr professionelles Verhalten zu verbessern (Gärtner & Vogt, 2013). Die Unterrichtsqualität von Lehrkräften, die in einem niederländischen Projekt Feedback von Schüler:innen zum Unterricht erhielten, verbesserte sich im Laufe der Zeit zwar leicht, aber nicht nachhaltig (Bijlma, Visscher, Dobbelaer & Veldkamp, 2019). Möglicherweise lag dies daran, dass die Lehrkräfte zwar Feedback erhielten, aber nicht systematisch bezüglich der Reflexion und Verbesserung der Unterrichtsqualität unterstützt wurden. Daher könnte die Besprechung der Feedback-Ergebnisse und die gemeinsame Reflexion von Handlungsalternativen für das Lehrverhalten dazu beitragen, den Unterricht von Lehramtsstudierenden zu verbessern.

Zur Unterrichtsqualität existieren unterschiedliche Konzepte, z.B. die drei Kerndimensionen konstruktive Unterstützung, effektives Klassenmanagement und kognitive Aktivierung (Klieme et al., 2009), 12 Ansätze zu inhaltsübergreifenden, mathematikspezifischen und hybriden Rahmenmodellen (Praetorius & Charalambous, 2018) oder drei Basisdimensionen der Unterrichtsqualität: Unterstützung des Wissenserwerbs, Motivation und Klassenraummanagement (Gärtner, Thiel, Nachbauer & Kellermann, 2021). Bisher existieren nur wenige Instrumente zur Unterrichtsqualität, die sich auf einzelne Unterrichtseinheiten beziehen, zeiteffizient am Ende des Unterrichts erhoben werden können und für Nachbesprechungen nützlich sind. Basierend auf den letztgenannten drei Basisdimensionen der Unterrichtsqualität beschreibt der vorliegende Beitrag daher die Entwicklung, Validierung und Erprobung eines Kurzfragebogens für Schüler:innen zur Rückmeldung der Unterrichtsqualität.

Zunächst wurden Items zur Unterrichtsqualität entwickelt, gemeinsam mit Expert:innen reduziert und durch erfahrene Lehrkräfte auf ihre Verständlichkeit geprüft und adaptiert. Studie 1 (N=77 Schüler:innen, 6 Unterrichtseinheiten) pilotierte 29 Items in verschiedenen Fächer wie Mathematik, Sport, Naturwissenschaften und Sozialwissenschaften. Auf der Grundlage von Item- und Skalenanalysen und der Korrelation mit einer Gesamtbewertung der Unterrichtsqualität wurden 15 Items für eine Kurzskala zur Unterrichtsqualität für Studie 3 ausgewählt. Studie 2 (N=178 Schüler:innen, 10 Unterrichtseinheiten) erfasste Unterrichtsqualität mit den 29 Items sowie dem SEEQ (Marsh, 2007; Dresel & Grassinger, 2012) und einem Beobachtungsbogen zur Unterrichtsqualität (Thiel et al., 2020; Gärtner et al., 2021, 2022). Die Analysen verwiesen auf konvergente Validität und divergierende Perspektiven von Schüler:innen und Beobachter:innen.

Studie 3 (N=438 Schüler:innen, 33 Unterrichtseinheiten) untersuchte im Praxissemester die 15 Items der Kurzskala im Unterricht von Lehramtsstudierenden. In einem Kurzbericht wurden Ergebnisse rückgemeldet, damit Universitätsdozierende und Lehramtsstudierende die höchsten oder niedrigsten Mittelwerte der Items mit den Stärken und Schwächen des Unterrichts vergleichen und Dimensionen der Unterrichtsqualität reflektieren konnten.

In vier Gruppendiskussionen mit N=17 Lehramtsstudierenden (3-5 Teilnehmer pro Gruppe) zeigten sich eine positive Wahrnehmung des Schüler:innen-Feedbacks, teils divergierende Einschätzungen aus den unterschiedlichen Perspektiven, eine Hinterfragung von zu „positivem“ Feedback sowie die Nützlichkeit der Kommentierung des Schüler:innen-Feedbacks durch die Universitätsdozierenden. Insgesamt scheint die Einbettung der Ergebnisse in ein Unterrichtsnachgespräch hilfreich, um deren

Kontextualisierung und Interpretation zu erleichtern und die Reflexion über das Unterrichtshandeln zu fördern. Die Ergebnisse deuten auf den Nutzen eines Feedback-Tools für Schüler:innen zur Unterrichtsqualität hin. Allerdings sprechen die reinen Daten nicht für sich selbst, sondern sind hilfreicher, wenn sie zusammen mit Lehrexpert:innen interpretiert werden. Das Feedback zu einer Unterrichtsstunde hat das Potenzial, das produktive Lernen von Lehramtsstudenten zu unterstützen.

Wie steht es um die angebotsseitige Qualität von Unterrichtsstunden Studierender im Praxissemester? Befunde und Implikationen einer Videostudie zum Literaturunterricht

Florian Hesse
Friedrich-Schiller-Universität Jena, Germany

In den letzten Jahren wurden sowohl in Deutschland als auch international an zahlreichen Hochschulstandorten verlängerte Praxisphasen im Lehramtsstudium implementiert (Darling-Hammond 2017; Gröschner und Hascher 2019). Dem Ausbau der Praktika steht dabei ein Defizit an Studien gegenüber, die professionelle Kompetenzen angehender Lehrpersonen nicht nur mittels Selbsteinschätzung, sondern mittels objektiver Verfahren erfassen. Ein Desiderat bilden dabei insbesondere Studien, die sich unterrichtlichen Lehr-Lernprozessen im Praktikum sowie deren Qualität widmen und dabei Aspekte von Fachlichkeit berücksichtigen (Ulrich et al. 2020; Wieser 2019). Ursächlich für diese Befundlage dürfte u.a. der bisherige Mangel an fachdidaktischen Arbeiten sein, die Unterrichtsqualität fachspezifisch modellieren und operationalisieren. In der Literaturdidaktik wurde diese Forschungslücke aufgegriffen, indem aktuelle generische Modelle von Unterrichtsqualität (z.B. Charalambous und Praetorius 2020) angebotsseitig mit Blick auf das literarische Lernen von Schüler:innen spezifiziert wurden. Dabei wird angenommen, dass sich fachliche Unterrichtsqualität einerseits in den Dimensionen „kognitiv-emotionale Aktivierung“ und „Auswahl und Aufbereitung von Lerngegenständen“ konkretisieren lässt und andererseits durch die fachübergreifend relevanten Qualitätsdimensionen „Klassenführung“ und „sozio-emotionale Unterstützung“ zu rahmen ist (Hesse und Winkler 2022).

Im Projekt „Literaturunterricht im Praktikum“ stellte sich vor diesem Hintergrund die Frage, inwiefern angehende Lehrpersonen des Faches Deutsch Lernangebote im Literaturunterricht qualitätsvoll gestalten. Aus einer Professionalisierungsperspektive heraus wurde dabei nicht nur untersucht, wie *die* Studierenden im Mittel hinsichtlich der o.g. Qualitätsdimensionen abschneiden, sondern auch, inwiefern interindividuelle Unterschiede zwischen den beobachteten Lernangeboten bestehen, welche die Notwendigkeit einer stärker heterogenitätssensiblen Lernbegleitung im Praxissemester nahelegen würden.

Zur Überprüfung der Fragestellungen wurde eine Gelegenheitsstichprobe von 22 videografierten Literaturstunden von 22 Lehramtsstudierenden des Faches Deutsch (5./6. Fachsemester) aus dem Jenaer Praxissemester herangezogen, die im Rahmen des Projekts OVID-Prax entstanden ist (Gröschner et al. 2019). Die angebotsseitige Unterrichtsqualität wurde in den o.g. Dimensionen durch drei externe, geschulte Rater:innen (Autor + 2 fortgeschrittene Studierende des Faches Deutsch) mittels hoch-inferenter Ratings (Pauli 2014) auf Basis eines eigens entwickelten Ratingmanuals (23 Items) mit vierstufigen Skalen (1 – geringe/wenig adaptive Ausprägung; 4 – hohe/adaptive Ausprägung) eingeschätzt. Anschließend wurde die Inter-Rater-Reliabilität durch Berechnung von

Intra-Klassenkorrelationskoeffizienten (ICC) überprüft (Wirtz und Caspar 2002). Die Auswertung der Ratings erfolgte dann einerseits deskriptiv-statistisch, um allgemeine Qualitätstendenzen feststellen zu können, andererseits wurden agglomerative hierarchische Clusteranalysen genutzt (Bacher et al. 2010), um explorativ zu prüfen, inwiefern sich unterschiedliche Muster (Cluster) in der Angebotsgestaltung bezogen auf die einzelnen Qualitätsdimensionen identifizieren lassen.

Im Ergebnis zeigt sich in Übereinstimmung mit anderen Videostudien im Bereich des sprachlich-künstlerischen Unterrichts bei erfahrenen Lehrpersonen (z.B. Lotz 2016; Riegler et al. 2022) sowie zum Unterricht Studierender im naturwissenschaftlichen Unterricht (z.B. Szogs et al. 2016; Vogelsang 2014), dass insbesondere Aspekte fachlicher Unterrichtsqualität („kognitiv-emotionale Aktivierung“, „Auswahl und Aufbereitung von Lerngegenständen“) im Mittel die größten Entwicklungspotenziale aufweisen (*M* jeweils < 2.5). Generische Dimensionen von Unterrichtsqualität („Klassenführung“, „sozio-emotionale Unterstützung“) werden demgegenüber mit hoher Qualität realisiert (*M* jeweils > 3.0), wobei hier erhebungsbedingte Verzerrungseffekte nicht unterschätzt werden dürfen (z. B. Selbstselektion der Klassen durch die Studierenden, Kameraeffekte). Die Clusteranalysen verdeutlichen zudem, dass hinsichtlich der Realisierung der einzelnen Qualitätsdimensionen unter den Studierenden eine enorme Spannbreite vorliegt. So konnten bspw. in der Dimension kognitiv-emotionale Aktivierung vier Cluster identifiziert werden, die von einem durchweg gering kognitiv-emotional aktivierenden Unterricht bis hin zu unterschiedlichen Varianten eines in hohem Maße kognitiv-emotional aktivierenden Unterrichts reichen. Die Befunde werfen damit weiterführende Fragen in zwei unterschiedliche Richtungen auf: Erstens ist in geplanten *Anschlussuntersuchungen* der Frage nachzugehen, welche individuellen Dispositionen (z.B. Wissensbestände, Überzeugungen, Wahrnehmungsfähigkeiten) oder Kontextbedingungen (z.B. Qualität des Mentorings) für die aufgezeigten Qualitätsdifferenzen entscheidend sind. Zweitens und darauf aufbauend ließe sich hinsichtlich der *Gestaltung von Lernbegleitungsformaten* im Praktikum fragen, ob und inwiefern diese zukünftig stärker als bisher mit der Heterogenität der Studierenden umgehen können und sollten. Der Vortrag stellt diese Fragen ausgehend von einer Präsentation zentraler Befunde zur Diskussion.

Verallgemeinerndes Potenzial in schulischen Aufgaben erkennen und nutzen – Die Sicht der Lehrkräfte

Birgit Vogt, **Markus Linnemann**
Universität Koblenz, Germany

Sprachliche Bildung stellt für Schüler:innen mit und ohne sonderpädagogischem Förderbedarf eine der Schlüsselqualifikationen in unserer hochliteralisierten Gesellschaft dar und ist zugleich eine für den Unterricht relevante Diversitätslinie. Personen ohne ausreichende sprachliche Bildung sind zunehmend bedroht, von gesellschaftlichen Diskursen ausgeschlossen zu werden (z.B. Gogolin, 2010). Das Verstehen von Narrativen und eine aktive Beteiligung an Diskursen, wie sie in vielfältiger Weise in unserer Zeit geführt werden (z.B. Foucault, 1970), setzt in den meisten Fällen komplexe und abstrakte Kenntnisse (z.B. zu „Klimawandel“, „Pandemie“ oder „Annexion“) voraus, bei denen verschiedene Perspektiven und Interpretationen integriert werden müssen. Im schulischen Kontext kommt somit dem Generalisieren eine wichtige epistemische Funktion zu (KMK, 2004; KMK, 2022). Die Schaffung von Zugängen zur Kompetenz des Verallgemeinerns für alle Schüler:innen ist vor dem

Hintergrund einer heterogenen Schüler:innenschaft in inklusiven Klassen besonders relevant. In der deutschsprachigen Bildungs- und Unterrichtsforschung wird dem Generalisieren jedoch bisher nur wenig Aufmerksamkeit geschenkt.

Im Kontext bildungssprachlicher Funktionen stellt das Generalisieren nach Feilke (2012) eine Voraussetzung für Erkenntnis und Erkenntniskommunikation dar. Es ist somit eine für die Unterrichtsforschung zentrale Größe. Verallgemeinerungsprozesse, wie Analysieren, Evaluieren und Kreieren in der Bloomschen Lernzieltaxonomie (Anderson & Krathwohl, 2001), sind Kern eines kognitiv aktivierenden Unterrichts. Eine eng mit dem Generalisieren verknüpfte, möglicherweise konstitutive, Operation ist das Bilden von Kategorien. Kategorisierungen stellen einen zentralen Aspekt der menschlichen Informationsverarbeitung dar, kognitive Kategorien sind „Produkte der Interaktion von Strukturen in der Welt und informationsverarbeitenden Systemen“ (Müsseler & Rieger, 2017). Wir verstehen Generalisieren als einen sprachlich-kognitiven Prozess der Abstraktion von einem konkreten, im Äußerungskontext eindeutig identifizierbaren Referenzobjekt (dem Exemplar) hin zu einer übergeordneten Kategorie bzw. einer übergeordneten Konzeption, die alle Referenzobjekte (Exemplare) mit gemeinsamen Merkmalen zusammenfasst. Beim umgekehrten Prozess der Spezifizierung bzw. beim Transfer wird eine Kategorie auf ein (anderes) Exemplar, das adäquate Merkmale aufweist, angewandt. Eine mögliche, an das Generalisieren heranführende kognitive Operation ist das Vergleichen.

Beide Prozesse – Generalisieren und Spezifizieren – spielen im Unterricht eine herausgestellte Rolle. Somit muss von Lehrkräften bei Aufgabenstellungen im Rahmen der Unterrichtsplanung das generalisierende bzw. spezifizierende Potenzial adressiert werden. Im Folgenden liegt der Fokus auf den Lehrkräften, die die Zielkompetenz des Generalisierens vermitteln. Denn nur, wer dies bei der Planung im Blick behält, ist in der Lage, den Unterricht dahingehend auszurichten, ihn sprachlich adäquat zu gestalten und insbesondere in heterogenen Schulklassen alle Schüler:innen bestmöglich zu fördern. Hieraus leitet sich die Fragestellung ab, ob von Lehrkräften in Aufgaben inhärentes Potenzial zum Generalisieren erkannt wird und wie es im Zuge der inklusiven Unterrichtsplanung berücksichtigt wird.

Der Fragestellung wird zurzeit mit Hilfe einer Online-Befragung von Lehrkräften aus neun Bundesländern nachgegangen. Kern der Befragung ist die Darstellung verschiedener Aufgabenstellungen aus Lehrwerken des Sachunterrichts der Grundschule, die mehr oder minder großes Potenzial des Generalisierens für den gewählten Themenbereich besitzen. Erste Ergebnisse machen deutlich, dass ein bestehendes Potential für die Förderung des Generalisierens als Zielkompetenz nur in geringem Maße gesehen wird. Lediglich ca. 14% der Lehrkräfte erkannten das Generalisierungspotenzial von hierzu geeigneten Aufgaben; ca. 18% gaben an, übergeordnete Konzepte mit Hilfe der dargebotenen Aufgaben vermitteln zu können. Immerhin erkannten 58% eine Eignung der entsprechenden Aufgaben zum Vergleichen als eine mögliche kognitive Operation für das Generalisieren. 18% der Lehrkräfte gaben an, sich für das Generalisieren anbietende Aufgaben Lernenden mit Förderbedarf vorzuenthalten, wohingegen sie sich weniger eignende Aufgaben darbieten würden. Es zeigten sich keine Unterschiede zwischen den Lehrkräften an Schulen mit oder ohne inklusivem Schwerpunkt.

Insgesamt zeigte sich, dass insbesondere ein stärkeres Bewusstsein geschaffen werden muss, Aufgabenstellungen zur Generalisierung zu nutzen und so z.B. im Sachunterricht Referenzobjekte tatsächlich „exemplarisch“ zu betrachten, d.h. mit der Zielrichtung Generalisieren. Im Vortrag werden Konstrukt, Studie und Ergebnisse dargelegt und vor diesem Hintergrund insbesondere im Kontext didaktischen Professionswissens zu inklusiver Beschulung diskutiert.

D-10: Motivation und Interesse

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S05 R03 H20
Chair der Sitzung: Stanislaw Schukajlow

Situationales Interesse beim Lösen offener und geschlossener realitätsbezogener Aufgaben – die Bedeutung von individuellen Merkmalen und Kontextmerkmalen

Stanislaw Schukajlow[1], Janina Krawitz[1], Katrin Rakoczy[2]
[1]Universität Münster, Institut für Didaktik der Mathematik und der Informatik; [2]Universität Gießen, Institut für Kindheits- und Schulpädagogik

Theorie

Interesse ist ein motivationales Konstrukt, welches eine Beziehung zwischen einer Person und dem Interessensgegenstand beschreibt und positiv mit Leistungen zusammenhängt (Krapp, 2005; Renninger & Hidi, 2016; Schiefele et al., 1993). Es wird zwischen einem zeitlich stabilen individuellen Interesse und Interesse in einer spezifische Situation (situationalem Interesse) unterschieden (Hidi & Renninger, 2006). Es wird angenommen, dass Aufgabenmerkmale, individuelle Merkmale und Kontextmerkmale für das Entstehen von situationalem Interesse wichtig sind. Es ist von Bedeutung, inwieweit sich das Interesse an offenen und geschlossenen Aufgaben (Aufgabenmerkmale) voneinander unterscheidet, ob dieser Unterschied in verschiedenen Schulformen (Kontextmerkmale; Realschule vs. Gymnasium) vorliegt und inwieweit situationales Interesse von individuellem Interesse in Mathematik und von der Leistung (individuelle Merkmale) vorhergesagt werden kann. Hierbei werden unter offenen (nicht wohl-definierten) Problemen realitätsbezogene Aufgaben verstanden (Modellierungsaufgaben), in denen einige Informationen fehlen und während der Bearbeitung beschafft oder geschätzt werden müssen (Jonassen, 2000; Silver, 1995; Yeo, 2017). Bei geschlossenen Aufgaben sind alle Informationen vorgegeben. Während in der Schule oft geschlossene Aufgaben im Mathematikunterricht bearbeitet werden, kommen in der Realität oft offene Aufgaben vor (Niss & Blum, 2020). Die Erforschung von Interesse an offenen Aufgaben stellt somit ein wichtiges Ziel der fachdidaktischen Forschung dar.

Hypothesen

1. Aufgrund der größeren Nähe offener Aufgaben zur Realität (Niss & Blum, 2020), erwarten wir höheres Interesse an offenen Aufgaben als an geschlossenen Aufgaben bei Lernenden in Realschule und Gymnasium.
2. Auf der Basis von Interessenstheorien (Krapp, 2005; Renninger & Hidi, 2016) erwarten wir, dass individuelles Interesse und Leistungen das Interesse an offenen und geschlossenen Aufgaben bei Lernenden in Realschule und Gymnasium vorhersagen können.

Methode

143 Neuntklässler:innen (51% weiblich, *M* = 16 Jahre) aus Realschulen (*N* = 76) und Gymnasien (*N* = 67) haben an einer Befragung zum individuellen Interesse an Mathematik teilgenommen und einen Leistungstest mit offenen und geschlossenen Aufgaben bearbeitet. Nach der Bearbeitung jeder Aufgabe berichteten sie sofort ihr situationales Interesse an der Bearbeitung dieser Aufgabe. Für die Interessenbefragungen wurden erprobte Skalen genutzt (1=stimmt gar nicht, 5=stimmt genau; Frenzel et al., 2012). Die interne Konsistenz der Skala zum individuellen Interesse war .84 (6 Items) und für Interesse an offenen und geschlossenen Aufgaben .78 und .83 (je 6 Items). Der Leistungstest bestand aus erprobten Aufgaben und die interne Konsistenz war für beide Aufgabentypen größer als .71 (je 6 Items).

Ergebnisse und Diskussion

Entgegen den Erwartungen zeigte eine ANOVA mit Innersubjektfaktor „Aufgabentyp“ (offen vs. geschlossen) und Zwischensubjektfaktor „Schulform“ (Realschule vs. Gymnasium) keine Unterschiede zwischen dem Interesse an beiden Aufgabentypen ($F(1, 137) = .232$; $p = .37$). Allerdings wurde eine Interaktion Aufgabentyp*Schulform beobachtet. Gymnasiast:innen tendierten zu höherem Interesse an geschlossenen Aufgaben ($M_{Gy_O} = 2.98$, $SD = 0.88$; $M_{Gy_G} = 3.12$, $SD = 0.93$; $t(66) = 1.825$, $p = .072$), während Realschüler:innen stärker an offenen Aufgaben interessiert waren ($M_{R_O} = 3.01$, $SD = 1.05$; $M_{R_G} = 2.79$, $SD = 1.05$; $t(71) = 2.905$, $p = .003$).

Die Pearson-Korrelationen zwischen individuellem Interesse und situationalem Interesse an beiden Aufgabentypen waren sehr hoch (.941 und .946). Die Korrelationen zwischen den Leistungen und Interesse an offenen Aufgaben waren nicht signifikant und zwischen den Leistungen und Interesse an geschlossenen Aufgaben waren niedrig, aber signifikant ($r_{Gy_O} = .372$, $r_{R_G} = .283$). Eine lineare Regression mit den Prädiktoren individuelles Interesse und Leistungen, zeigte dass individuelles Interesse aber nicht die Leistungen das situationale Interesse bei geschlossenen Aufgaben vorhersagen kann ($\beta_{Gy_O} = .94$, $\beta_{R_G} = .93$).

Die Ergebnisse bestätigen eine hohe Bedeutsamkeit von Aufgaben- und Kontextmerkmalen sowie von individuellen Merkmalen für das situationale Interesse. Individuelles Interesse an Mathematik zeigte sich als hochprädiktiv für die Bearbeitung von beiden Aufgabentypen. Eine praktische Implikation ist die Möglichkeit der Steigerung des situationalen Interesse in der Realschule durch die Behandlung offener Aufgaben.

Goals, behavior, success – the complex triadic association of achievement goals, e-learning behavior, and exam performance

Marc Philipp Janson, Stefan Janke
University of Mannheim, Germany

Intelligent tutoring systems that entail exercises and repetition material provide new ways to support students in their (self-regulated) learning process (Kulik & Fletcher, 2016). Yet little is known regarding *why* and *how* students use these tools and how differences in usage explain differences in effectiveness. In our striving to answer these open research questions, we draw from the achievement goal approach (Daumiller et al., 2019; Murayama et al., 2012) that suggests that different achievement goals (i.e., describing aspired end-states in achievement situations) facilitate different patterns of learning (Diseth & Kobbeltvedt, 2010). Considering the existing literature, we assume that learning and task approach goals motivate learners to use intelligent tutoring systems in a way that facilitates deeper learning (more invested learning time, which is distributed rather than massed, and less procrastination). In contrast, we assume that normative avoidance, appearance avoidance, and work avoidance goals are associated with patterns of surface learning (massed learning and procrastination). Additionally, we assumed that normative approach and appearance approach goals would be associated with massed learning. However, as individuals must weigh to which degree they invest time in learning activities, we assume that the actual use of intelligent tutoring systems also depends on whether individuals think that using intelligent tutoring systems assists them in achieving their respective goals. This perceived instrumentality of intelligent tutoring systems for goal striving is meant to moderate the postulated main effects of achievement goal content. As it is reasonable to

assume that deep processing (i.e., high learning times, which are distributed over a large time span; Soderstrom & Bjork, 2015), as well as low procrastination scores (Steel, 2007), facilitate optimal learning, we also assume that achievement goals are indirectly tied to later performance through their association with patterns of learning.

To investigate our research questions, we collected a) self-report data on learners' achievement goals as well as the perceived instrumentality of intelligent tutoring systems for goal striving at the beginning of a semester, b) process-data (i.e., logfiles) reflecting the usage of an intelligent tutoring system and as such patterns of learning (absolute learning time, distribution of learning time, procrastination/onset of learning) throughout the semester and c) exam performance at the end of semester. Hypotheses, operationalization of the main variables, and research design were preregistered in advance (https://aspredicted.org/4SY_B4Q). Our net sample included data from 92 university students using an intelligent tutoring system to prepare for a high-stakes statistics exam.

A series of latent moderated mediation models revealed mixed evidence for our hypotheses: We found that particularly task approach goals were linked to deep learning strategies such as the investment of learning time (β = 0.26, p = .016) as well as less massed learning (β = -0.15, p = .049). Both effects were moderated by the perceived instrumentality of the intelligent tutoring system for task goal striving (time: β = 0.19, p < .001; distribution: β = 0.19, p = .031) as predicted. We also found negative associations of appearance avoidance goals with distributed learning (β = -0.24, p = .015). We found no significant associations between procrastination and achievement goals. However, procrastination within the software was the only significant predictor for exam success (β = -0.45, p < .001).

In sum, (some) achievement goals are suitable to predict the way learners interact with intelligent tutoring systems but the main learning parameter that tied the usage of such tools to actual performance was not predicted by goal striving. Given that prior research often (solely) relied on self-reports, our research provides a more accurate (and partly challenging) picture of associations between achievement motivation, actual learning behavior, and performance.

Motivationale und emotionale Aspekte bei Open-Book Klausuren

Anne Schlosser, Jakob Meyer, Jennifer Paetsch
University of Bamberg, Germany

Theoretischer Hintergrund und Fragestellung

Closed-Book Prüfungen, in denen die Nutzung von Materialien in der Prüfungssituation untersagt oder beschränkt (z.B. auf Formelsammlungen) ist, sind die gängigste Prüfungsart im Bereich der Higher Education. Vorteile dieser Prüfungsform sind u.a. Effektivität, Standardisierung und Sicherheit gegen Täuschungsversuche. Als Einschränkungen werden beispielsweise negative Emotionen und wenig organisatorische Flexibilität aufgeführt. Ein alternatives Prüfungsformat sind Open-Book Prüfungen, in denen Studierende während der Prüfung auf Ressourcen zugreifen können. Dieses Prüfungsformat scheint nicht zuletzt aufgrund der Zunahme von Online-Lehrangeboten relevant (vgl. Parker et al., 2021).

Es existieren widersprüchliche empirische Befunde zum Vergleich der Leistungen in beiden Prüfungsformaten (z.B. Ghrabi et al., 2012; Rummer et al., 2019). In weiteren Studien wurde die Wirkung des zugelassenen Materials in Open-Book Prüfungen untersucht. So zeigen Niedermeier et al. (2022), dass die Anzahl der Unterlagen (nicht die Qualität), die in eine Open-Book-Prüfung mitgebracht werden, mit dem gezeigten konzeptuellen und prozeduralen Wissen zusammenhängen. Ghrabi et al. (2012) untersuchten Open-Book Prüfungen mit Cheat Sheets (eigenständig angefertigt, ohne inhaltlichen Vorgaben). Die Qualität der Cheat Sheets korrelierte mit der Prüfungsleistung. Weitere Studien zeigen eine verringerte Prüfungsangst in Open-Book Prüfungen (z.B. Riegel & Evans, 2021). Darüber hinaus existiert wenig Forschung zu emotional-motivationalen Aspekten bei Open-Book Prüfungen bzw. während deren Vorbereitung (z.B. Niedermeier et al., 2022).

Hier setzt der vorliegende Beitrag mit zwei Forschungsfragen an:

1) Wirkt sich die Nutzung der Booklets (handschriftliche Zusammenfassungen) positiv auf die prüfungsbezogene Selbstwirksamkeit und die Prüfungsleistung aus?

2) *Welche Rolle spielen dabei emotional-motivationale Aspekte?*

Basierend auf der Control-Value-Theorie (Pekrun, 2006) wird angenommen, dass Emotionen und Motivation mit der Prüfungsleistung zusammenhängen. Weiterhin sind Emotionen eine Quelle der Selbstwirksamkeitserwartung (Bandura, 1997). Basierend auf den Theorien und bisherigen Forschungsergebnissen wurde folgendes Forschungsdesign entwickelt:

In einer Psychologievorlesung konnten Lehramtsstudierende wöchentlich handschriftliche Zusammenfassungen des Vorlesungsinhalts (Booklets) abgeben, welche digitalisiert wurden und in der Abschlussprüfung zur Verfügung standen (Open-Book-Prüfung). Die Verwendung der Booklets war freiwillig. Die Abschlussprüfung beinhaltete noch einen zweiten Prüfungsteil zu einer zweiten Vorlesung, in der keine Booklets zugelassen waren (Closed-Book-Prüfung). Folgende Hypothesen werden überprüft:

Hypothese 1: Die Studierenden haben (a) eine höhere prüfungsbezogene Selbstwirksamkeit und (b) eine bessere Leistung in dem Prüfungsteil, den sie mit Booklets absolvieren dürfen, im Vergleich zum anderen Prüfungsteil (keine Booklets).

Spezifische Hypothesen für die Open-Book Prüfung:

Hypothese 2: Die Anzahl der abgegebenen Booklets wirkt sich positiv auf die Prüfungsleistung aus, unter Kontrolle der prüfungsbezogenen Selbstwirksamkeit und der subjektiven Bedeutung der Prüfung.

Hypothese 3: Motivation, Einstellung zum Booklet, emotional states und der eingeschätzte Nutzen des Booklets korrelieren mit (a) prüfungsbezogener Selbstwirksamkeit und mit (b) Leistung im Prüfungsteil mit Booklets.

Methode

Im Sommersemester 2022 nahmen 223 (67% weiblich) Lehramtsstudierende, die sich im Mittel im 2. Semester befanden, an der Studie teil, wobei 71 Studierende die Booklets nutzten. Zur Überprüfung der Fragestellung wurden die prüfungsbezogene Selbstwirksamkeit (adaptiert Schwarzer &

Jerusalem, 1995), emotionales Arousal und Valenz (Russel, Weiss & Mendelsohn, 1989), Motivation (Justus, 2017) und die erreichte Leistung im jeweiligen Prüfungsteil ermittelt. Als Kontrollvariablen wurden zusätzlich Einstellungen zum Booklet (adaptiert Laugwitz et al., 2008) sowie motivationale Faktoren (persönliche Bedeutungszuschreibung, aufgewendete Kosten, Nützlichkeit) erfasst.

Vorläufige Ergebnisse

Vorläufige Ergebnisse zeigen, dass die Studierenden unter Kontrolle der Prätestergebnisse ($t(150)=-.53$, $p=.299$) eine signifikant höhere Selbstwirksamkeit hinsichtlich des Prüfungsteils haben, den sie mit Booklets absolvieren können ($t(73)=4.98$, $p<.001$). Außerdem erreichen die Studierenden im Open-Book Prüfungsteil signifikant bessere Ergebnisse ($t(118)=11.92$, $p<.001$). Hypothese 1 kann somit vorläufig bestätigt werden. Eine weitere Datenerhebung, in der Booklets im anderen Prüfungsteil zugelassen werden, steht noch aus. Korrelationsanalysen zeigen zunächst signifikante Zusammenhänge von Selbstwirksamkeit mit Motivation ($r=.715$, $p<.001$) sowie mit Einstellungen zum Booklet (bspw. Kreativität, $r=.251$, $p=.036$), jedoch nicht mit Emotionen. Weitere Analysen stehen noch aus. Die Ergebnisse werden im Kontext der Control-Value-Theorie sowie der Selbstwirksamkeitstheorie diskutiert.

Feedback von Grundschullehrkräften im Zusammenhang mit den Selbstwirksamkeitsüberzeugungen und der Motivation von Kindern im naturwissenschaftlichen Sachunterricht

Fabian Hoya, Jan Roland Schulze, Eva Blumberg, Frank Hellmich
Universität Paderborn, Germany

Das Feedback von Lehrkräften stellt eine wichtige Gelingensbedingung für das Lernen von Kindern im naturwissenschaftlichen Sachunterricht der Grundschule dar. Feedback wird dabei als eine Information verstanden, die Lehrkräfte Kindern während ihrer Lernprozesse oder als Folge von erbrachten Leistungen senden (vgl. Hattie & Timperley, 2007). Positives und negatives Feedback kann – folgt man Bandura (1997) – die Selbstwirksamkeitsüberzeugungen von Kindern beeinflussen und sich entsprechend auf ihr Lernverhalten auswirken (vgl. auch Schunk & Pajares, 2009). Unter Selbstwirksamkeit wird dabei die Überzeugung eines Individuums verstanden, das für ein erwünschtes Ergebnis erforderliche Verhalten erfolgreich ausführen zu können. Kinder mit gering ausgeprägten Selbstwirksamkeitsüberzeugungen nehmen sich in herausfordernden Situationen häufig als hilflos wahr. In verschiedenen Studien konnten förderliche Effekte von positivem Feedback auf die Selbstwirksamkeitsüberzeugungen von Schüler*innen und ihr Lernverhalten – wie beispielsweise ihre Interessen an den Unterrichtsinhalten – nachgewiesen werden (Won, Lee & Bong, 2017; Xu, Du & Fan, 2017). Ebenso konnten auch motivationshemmende Effekte negativen Feedbacks auf die Lernprozesse von Schüler*innen aufgezeigt werden (Morgan, 2001). Gegenwärtig ist allerdings noch weitgehend ungeklärt, ob und inwiefern positives bzw. negatives Feedback die Selbstwirksamkeitsüberzeugungen von Kindern im Grundschulalter begünstigt bzw. beeinträchtigt und sich auf deren intrinsische Motivation für die Beschäftigung mit den Unterrichtsinhalten auswirkt. Unter intrinsischer Motivation wird dabei der Wunsch oder die Absicht“ verstanden, „eine bestimmte Handlung durchzuführen, weil die Handlung selbst als interessant, spannend, herausfordernd […] erscheint“ (Schiefele & Köller, 2010, S. 336).

Auf der Grundlage der theoretischen Annahmen von Bandura (1997) sowie Schunk und Pajares (2009) nehmen wir (1) an, dass positives Feedback prädiktiv für die Selbstwirksamkeitsüberzeugungen von Kindern und ihre intrinsische Motivation für die Beschäftigung mit Inhalten des naturwissenschaftlichen Sachunterrichts ist. Umgekehrt vermuten wir (2), dass negatives Feedback eng mit dem Hilflosigkeitserleben von Kindern im naturwissenschaftlichen Sachunterricht korreliert sowie negativ mit deren intrinsischer Motivation zusammenhängt.
An unserer Untersuchung waren N = 669 Kinder der dritten und vierten Jahrgangsstufe beteiligt. Die Kinder wurden im Rahmen eines Fragebogens darum gebeten, Stellung zu dem von ihnen wahrgenommenen positiven und negativen Feedback ihrer Lehrkräfte, ihren Selbstwirksamkeitsüberzeugungen, ihrem Hilflosigkeitsempfinden sowie zu ihrer intrinsischen Motivation im naturwissenschaftlichen Sachunterricht zu nehmen. Für die Überprüfung der hypothetisch angenommenen Zusammenhänge der einzelnen Variablen wurde ein Strukturgleichungsmodell in Mplus berechnet.
Die Ergebnisse unserer Untersuchung zeigen, dass (1) das von den Kindern wahrgenommene positive Feedback zwar ein signifikanter Prädiktor für ihre Selbstwirksamkeitsüberzeugungen, aber nicht für ihre intrinsische Motivation im naturwissenschaftlichen Sachunterricht ist. Weiterhin wird (2) deutlich, dass das Hilflosigkeitsempfinden der Kinder durch das perzipierte negative Feedback erklärt werden kann und auch negativ mit ihrer intrinsischen Motivation im naturwissenschaftlichen Sachunterricht korreliert.
Die von uns ermittelten Befunde geben damit Hinweise auf förderliche und hemmende Effekte von Feedback auf die Selbstwirksamkeitsüberzeugungen von Kindern und ihr Lernverhalten im naturwissenschaftlichen Sachunterricht der Grundschule. In weiterführenden Untersuchungen sollte der Frage nachgegangen werden, ob und inwiefern konkrete Maßnahmen – wie beispielsweise der Einsatz zielführenden positiven Feedbacks – die Selbstwirksamkeitsüberzeugungen und die Motivation von Kindern im naturwissenschaftlichen Sachunterricht begünstigen können.

Symposien

D-11: Effektivität von Rater*innentrainings: Wie gelingt die Beobachtung der Unterrichtsqualität durch externe Beobachtende?

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* S07 S00 D07

Effektivität von Rater*innentrainings: Wie gelingt die Beobachtung der Unterrichtsqualität durch externe Beobachtende?

Chair(s): **Julia Larissa Maier** (Institut für Bildungsanalysen Baden-Württemberg, Deutschland), **Evelin Ruth-Herbein** (Institut für Bildungsanalysen Baden-Württemberg, Deutschland)
*Diskutant*in(nen):* **Sandy Taut** (Bayerisches Landesamt für Schule)

Die Bedeutung der Unterrichtsqualität für die kognitive sowie emotional-motivationale Entwicklung von Schülerinnen und Schülern ist seit Jahrzehnten Gegenstand der empirischen Bildungsforschung (Hattie & Zierer, 2017; Helmke, 2017). Um Aussagen über die Qualität eines Unterrichts zu treffen ist ein häufig gewählter Zugang (sowohl in der Schulpraxis als auch in der Bildungsforschung) der Einsatz von Beobachtungsbögen, die von externen Rater*innen ausgefüllt werden (Clausen, 2002; Tarkian et al., 2019).

Für eine objektive, reliable und valide Einschätzung der Unterrichtsqualität sind allerdings kriteriengeleitete Beobachtungssysteme notwendig. Um Beobachtungsfehler zu reduzieren und Einschätzungen zu standardisieren werden externe Beobachtende in der Regel für ihren Einsatz in der Unterrichtsbeobachtung geschult (Praetorius et al., 2013). Diese Trainings variieren dabei allerdings beispielsweise in Dauer und methodischer Umsetzung. Systematische Forschung dazu, welche Komponenten wirksamen Rater*innentrainings zugrunde liegen und wie Schulungen für Rater*innen mit unterschiedlicher Expertise erfolgreich umgesetzt werden können sind selten.

Die Beitragenden des Symposiums stellen sich daher die Frage, welche Voraussetzungen und Unterstützungssysteme notwendig sind, damit externe Beobachtende zu objektiven, reliablen und validen Ratings der Unterrichtsqualität kommen. Anhand empirischer und konzeptioneller Arbeiten soll herausgearbeitet werden, was ein effektives Rater*innentraining auszeichnet und ob und wie deren Effektivität überprüft werden kann. Gleichzeitig werden Fragen der Implementation der Trainings, welche häufig mit hohen zeitlichen Ressourcen verknüpft sind, adressiert.

Der erste Beitrag liefert zunächst einen Literaturüberblick zur Effektivität von Rater*innentrainings zur Unterrichtsqualität. Anhand einer Literaturrecherche wird der aktuelle empirische Forschungsstand dargestellt und herausgearbeitet, welche Hinweise es auf die Effektivität und die Gestaltung von Rater*innentrainings im Bereich der Bewertung von Unterrichtsqualität gibt.

Nach diesem Literaturüberblick werden im zweiten Beitrag Ergebnisse einer randomisiert kontrollierten Studie vorgestellt, die die Wirksamkeit und Durchführbarkeit zweier Qualifizierungsveranstaltungen im Kontext der Lehrkräfteaus- und -fortbildung untersucht und vergleicht. Die Qualifizierungen unterscheiden sich systematisch hinsichtlich zugrundeliegender Komponenten der Umsetzung.

Der dritte Beitrag gibt Einblicke in die Schulung von Fachpersonen für die Schulevaluation und thematisiert die Frage, ob und wie eine Zertifizierung im Rahmen der Schulung zur Qualitätssicherung beitragen kann. Hierfür werden die Zertifizierungsergebnisse der teilnehmenden Personen vorgestellt.

Anhand deren Übereinstimmung zu Masterratings können Rückschlüsse auf die Effektivität der Schulungsmaßnahme gezogen werden.

Die drei Beiträge werden anschließend mit Blick auf forschungs- und praxisrelevante Aspekte diskutiert. Das Symposium möchte damit den Austausch zu den Herausforderungen der Unterrichtsbeobachtung und den Umgang mit ebendiesen Herausforderungen befördern. Auf diese Weise sollen wichtige Erkenntnisse zur Effektivität und wirksamen Gestaltung von Rater*innentrainings zusammengetragen und diskutiert werden, die für die Bildungsforschung und -praxis von zentraler Bedeutung sind.

Beiträge des Symposiums

Effektivität von Rater*innentrainings zur Beurteilung von Unterrichtsqualität: Ein Literaturüberblick

Valerie Gitzi[1], Wida Wemmer-Rogh[1], Thilo Kleickmann[2], Mirjam Steffensky[3], Aiso Heinze[4], Anna-Katharina Praetorius[1]
[1]Universität Zürich, [2]Christian-Albrechts-Universität zu Kiel, [3]Universität Hamburg, [4]Leibniz-Institut für die Pädagogik der Naturwissenschaften

Theoretischer Hintergrund

Die Beurteilung von Unterricht mittels externer Beobachtung wird oft als Goldstandard der Unterrichtsbeurteilung angesehen (Helmke, 2017). Um die Beurteilung zu standardisieren und somit auch Beobachter*innenfehler zu reduzieren, erfolgt in der Regel in der Unterrichtsforschung, aber auch in angewandten Settings wie der Schulinspektion, ein Rater*innentraining (Praetorius, 2013). Dieses Training zielt im Kern auf eine Vereinheitlichung des Verständnisses von Unterrichtsqualitätsdimensionen des eingesetzten Instruments und der dazugehörigen Subdimensionen, Items und Indikatoren ab (Joe et al., 2013). Dadurch soll die Reliabilität der Beurteilungen verbessert werden. Allerdings fehlen bislang systematische Erkenntnisse zur Effektivität der Trainings und dazu, wie Trainings gestaltet werden müssen um effektiv zu sein (Lemay et al., 2021).

Fragestellung

Entsprechend sollen mittels eines Literaturreviews die folgenden Fragen bearbeitet werden:

(1) Wie effektiv sind Rater*innentrainings zur Beurteilung von Unterrichtsqualität im Hinblick auf die Verbesserung der Reliabilität und Validität der Beurteilungen?

(2) Welche Hinweise finden sich in der Literatur zur Gestaltung von Rater*innentrainings?

Methode

Um diese Fragen zu adressieren, wurde eine systematische Literaturrecherche mithilfe der Datenbanken ERIC, EBSCOHOST, Google Scholar, JSTOR, PSYINDEX und FIS-Pädagogik vorgenommen. Die Literatur spiegelt den Stand bis Sommer 2022 wider. Zudem wurden nach dem Schneeballprinzip weitere Publikationen ergänzt. In die Recherche wurden Trainings in Bezug auf 20 generische und fachspezifische Frameworks, z.B. CLASS (Classroom Assessing Scoring System) bzw. MQI (Mathematical Quality of Instruction) einbezogen. Insgesamt wurden so 56 Publikationen identifiziert. Angaben zur Effektivität und zu Gestaltungsmerkmalen der Trainings wurden extrahiert und mittels eines deduktiv-induktiven Vorgehens kategorisiert.

Ergebnisse und Diskussion

In Bezug auf die Effektivität von Rater*innentrainings zeigte sich, dass in vielen Publikationen keine Angaben zur Effektivität gemacht werden (z.B. Blazar et al., 2018; McCaffrey et al., 2015). Die existierenden Angaben zur Effektivität beziehen sich auf unterschiedliche Ansätze zur Bestimmung von Effektivität. Diese reichen von Analysen der Übereinstimmung der Rater*innen während (z.B. Lenske et al., 2015) oder nach (z.B. Broderick et al., 2019) dem Training bis hin zu einer formalen Zertifizierung (z.B. Hill et al., 2012).

Die angegebenen Übereinstimmungswerte variieren deutlich und reichen von 14 Prozent (Bill & Melinda Gates Foundation, 2012) bis 87 Prozent (Kelcey & Carlisle, 2013). Wird eine Zertifizierung vergeben und am Ende der Schulung die Übereinstimmung mit einem Masterrating berechnet, so reichen die Werte von 34 Prozent (vgl. Casabianca et al., 2015) bis 80 Prozent (vgl. Sailors et al., 2014).

In Bezug auf die Gestaltungsmerkmale von Rater*innentrainings lässt sich zunächst festhalten, dass einige Gestaltungselemente in nahezu allen einbezogenen Trainings vorkommen. So sind in den meisten Trainings Theorieteile und Übungsbeispiele (meist videographierte Unterrichtslektionen bzw. Ausschnitte) enthalten. (z.B. Silverman & Crandell, 2010; van der Lans et al., 2016)

Die ersten Analysen geben Hinweise darauf, dass sich die Dauer der Trainings sowie die Art und Weise des Einsatzes der Übungsbeispiele deutlich unterscheiden. Bei der Dauer der Trainings variieren die Werte zwischen zwei Stunden (vgl. Sherman & Camilli, 2014) und 180 Stunden (vgl. Nowinska & Praetorius, 2017).

Beim Einsatz der Übungsbeispiele wird in vielen Trainings auf eine an die individuellen Einschätzungen anschließende Diskussion fokussiert, in welchen Unklarheiten diskursiv gelöst werden (z.B. Martin-Raugh et al., 2016). Andere Trainings setzen hingegen vornehmlich auf eine Rückmeldung zur Qualität durch Abgleich mit Masterratings (z.B. Mantzicopoulos et al., 2018).

Bereits diese Ergebnisse liefern Hinweise dafür, dass die Effektivität und Gestaltung von Rater*innentrainings stark variieren. In weiteren vertiefenden Analysen sollen nach Möglichkeit Effektivität und Gestaltungsmerkmale in Beziehung gesetzt werden. Grundsätzlich zeigt die Recherche jedoch, dass methodische Optimierungen in Zukunft nötig sind, um verlässlichere Aussagen zur Effektivität der Trainings sowie zur Gestaltung effektiver Trainings tätigen zu können. Dazu gehören unter anderem der Einbezug von Kontrollgruppen und längsschnittliche Designs. Zudem sollte neben der Reliabilität die Validität der Beurteilungen als Effektivitätskriterium hinzugezogen werden.

Wie viel Austausch ist notwendig? Die Effektivität synchroner Austauschphasen bei Online-Schulungen zum Training der Be-obachtungsgenauigkeit

Julia Larissa Maier[1], Evelin Ruth-Herbein[1], Benjamin Fauth[2]
[1]Institut für Bildungsanalysen Baden-Württemberg, [2]Institut für Bildungsanalysen Baden-Württemberg; Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

Theoretischer Hintergrund

Als Instrument zur Unterrichts- und Schulentwicklung wurde am Institut für Bildungsanalysen Baden-Württemberg (IBBW) der Unterrichtsfeedbackbogen Tiefenstrukturen (kurz: UFB; Fauth et al., 2022)

entwickelt. Ziel ist es, die unterrichtlichen Tiefenstrukturen greifbar und beschreibbar zu machen und dadurch für alle Schularten, Klassenstufen und Unterrichtsfächer eine gemeinsame Grundlage für die Unterrichtsbeobachtung zu schaffen. Die psychometrischen Eigenschaften und die Einsetzbarkeit des Instrumentariums wurden bereits in mehreren Begleitstudien überprüft (Maier et al., 2021; Maier et al., in Druck).

Um Lehrkräfte im Umgang mit dem Instrumentarium zu schulen, wurde vom Zentrum für Schulqualität und Lehrerbildung Baden-Württemberg (ZSL) basierend auf aktuellen Erkenntnissen zur Gestaltung von Lehrkräftefortbildungen (vgl. Lipowsky & Rzejak, 2021) ein Qualifizierungskurs konzipiert. Dieser ist mit 20 Zeitstunden allerdings sehr zeitaufwändig. Um Beobachtungstrainings gut in die Praxis implementieren zu können, müssen diese möglichst zeit- und ressourcensparend und damit praktikabel sein. In der aktuellen Forschung gibt es allerdings kaum systematische Studien, die Erkenntnisse zur zeit- und ressourcensparenden Gestaltung von Rater*innentrainings und deren Effektivität liefern. Daher wurde der bereits erprobte und evaluierte Qualifizierungskurs systematisch adaptiert: mehr Selbstlern- und weniger Austauschphasen führen zu einem zeitlich verkürzten Qualifizierungskurs und einer höheren Flexibilität in der Durchführung, allerdings zu weniger Austausch- und Diskussionsmöglichkeiten. Im Vortrag werden die Ergebnisse einer randomisiert kontrollierten Wirksamkeitsstudie vorgestellt.

Fragestellung

Die zwei Durchführungsarten des Qualifizierungskurses (Austausch- und Selbstlerngruppe) werden vergleichend gegenübergestellt, um daraus Implikationen für die zukünftige Gestaltung von Rater*innentrainings abzuleiten.

1. Führt das Durchlaufen der Qualifizierungskurse zur Erreichung der intendierten Ziele?

a. Einstellung zum Instrumentarium

b. Sicherheit im Umgang mit dem Instrumentarium

c. reliable und valide Einschätzungen der Unterrichtsqualität

2. Unterscheiden sich die Teilnehmenden der Austausch- und der Selbstlerngruppe systematisch auf diesen Zielvariablen?

Methode

Der Qualifizierungskurs besteht aus drei Bausteinen, wobei sich die beiden Durchführungsarten dahingehend unterscheiden, ob synchrone Bearbeitungsphasen sowie Diskussionsmöglichkeiten in den Ablauf eingebettet sind: Teilnehmende der Austausch-Gruppe durchlaufen pro Qualifizierungsbaustein eine Inputphase im Selbstlernen. Im Anschluss an jede Inputphase findet ein synchrones Online-Austauschtreffen statt, in welchem gemeinsam eine Trainingsphase bearbeitet und das Erlernte reflektiert und diskutiert wird. Damit umfasst die Qualifizierung ca. 20 Zeitstunden. Teilnehmende der Selbstlern-Gruppe durchlaufen alle drei Bausteine asynchron und vollständig im Selbstlernen ohne dazwischengeschaltete Austauschtreffen. Damit umfasst die Qualifizierung für sie nur ca. 13 Zeitstunden.

Der Qualifizierungskurs fand von Ende April bis Anfang Juni 2022 in einem Zeitraum von sechs Wochen statt. An der Wirksamkeitsstudie nahmen 92 Lehrkräfte teil, welche in der Aus- und/oder Fortbildung für unterschiedliche Fachbereiche tätig sind (Nweiblich = 61; Nmännlich = 23; Berufserfahrung: M = 19.08 Jahre, SD = 7.5). Die Teilnehmenden wurden randomisiert einer der beiden Durchführungsarten zugeteilt.

Vor, zwischen und nach dem Qualifizierungskurs fanden Online-Befragungen statt. Mithilfe des Community-of-Inquiry-Frameworks (Garrison et al., 1999; Fragebogen adaptiert nach Ammenwerth & Hackl, 2017) wird untersucht, wie die Teilnehmenden den Qualifizierungskurs wahrnehmen und ob/in welchem Ausmaß sich die wahrgenommene Lehrpräsenz, kognitive Präsenz und soziale Präsenz zwischen den beiden Durchführungsarten unterscheidet. Zudem werden beispielsweise die Einstellung zum Instrumentarium und zur kollegialen Hospitation, die Selbstwirksamkeit bei der Unterrichtsbeobachtung sowie das selbsteingeschätzte Wissen zu Tiefenstrukturen erhoben. Darüber hinaus schätzen die Teilnehmenden vor und nach Durchlaufen des Qualifizierungskurses 10-minütige Videosequenzen der TALIS-Videostudie Deutschland (Grünkorn et al., 2020) hinsichtlich ihrer Unterrichtsqualität ein. Auf dieser Grundlage wird die Beobachtungsübereinstimmung vor und nach Durchlaufen des Qualifizierungskurses untersucht.

Ergebnisse und ihre Bedeutung

Die Datenerhebung lief bis August 2022, weshalb die Datenauswertung zum aktuellen Zeitpunkt noch nicht abgeschlossen ist. Bei der Tagung im März 2023 sollen erste Ergebnisse vorgestellt werden. Die Studie soll aufzeigen, welchen Beitrag die zum Teil sehr zeitintensiven Austausch- und Diskussionsphasen zur Wirksamkeit von Rater*innentrainings leisten. Hieraus können relevante Schlussfolgerungen sowohl für die Bildungsforschung als auch die Bildungspraxis gezogen werden, wie Schulungen zukünftig zeit- und ressourcenschonend aber dennoch wirksam gestaltet werden können.

Ist bestanden gut genug? Datenbasierte Reflexion von Möglichkeiten zur Qualifizierung in der beobachtungsbasierten Beurteilung von Unterrichtsqualität

Wida Wemmer-Rogh, Leonie Gossner, Anna-Katharina Praetorius
Universität Zürich

Theoretischer Hintergrund

In der Unterrichtsforschung stellen Unterrichtsbeobachtungen einen prominenten Ansatz für die Beurteilung von Unterrichtsqualität dar (Helmke, 2015). Die Bedeutung dieser Vorgehensweise zeigt sich auch darin, dass sich zunehmend politische Entscheidungsorgane, Schulbehörden oder Schulevaluationsfachstellen für die beobachtungsbasierte Beurteilung von Unterrichtsqualität interessieren (Taut & Rakoczy, 2016). Beobachtungen können allerdings mit Verzerrungstendenzen bei der Informationsaufnahme und -verarbeitung (Bell et al., 2018; Casabianca et al., 2015; Hill et al., 2012; Hoyt, 2000; Praetorius, 2014; Pietsch & Tosana, 2008) einhergehen. Studien betonen den zentralen Stellenwert von Beobachtungstrainings zur Reduzierung solcher Urteilsfehler (Lotz, et al., 2013; McClelland et al., 2012; Woehr & Huffcutt, 1994). Während einzelne Untersuchungen zeigen, dass Beobachtungstrainings zur Disziplinierung kognitiver Prozesse und zur Reliabilität und Validität von Unterrichtsbeurteilungen beitragen (Liu et al., 2019; Wang, 2010), lassen andere Studien, bei

denen Unterrichtsbeurteilungen trotz Trainings nur mäßig genau ausfallen (Bell et al. 2018; Praetorius, 2014), die Konzeption von Beobachtungstrainings hinterfragen. Eine Literaturrecherche existierender Trainings zur Einschätzung von Unterrichtsqualität (Gitzi et al., i. V.) zeigt, dass obwohl Beobachtungstrainings als Instrument zur Sicherstellung der Beurteilungsgüte gelten, keine einheitlichen Vorgehensweisen in der Trainingskonzeption existieren. Insbesondere hinsichtlich der Maßnahmen zur Qualitätssicherung wie Zertifizierungen, Kalibrierungen oder Validierungen gibt es unterschiedliche Handhabungen. Gleichzeitig finden sich wenige Auseinandersetzungen damit, welche Ansätze die Qualität von Unterrichtsbeurteilungen sichern. Zudem scheinen solche Qualifizierungsmaßnahmen im deutschsprachigen Raum bislang wenig etabliert, während insbesondere die Qualifizierung über Zertifizierung in den USA verbreitet ist (ebd.).

Fragestellung

Für eine stärkere Etablierung im deutschsprachigen Raum stellt sich die Frage, inwiefern eine Qualifizierung über Zertifizierung Anhaltspunkte dazu liefert, ob die Anwendung eines Beurteilungsinstrumentariums für Unterrichtsqualität angemessen trainiert wurde?

Methode

Dieser Frage wurde im Projekt „Qualifizierung in der Unterrichtsbeurteilung“ in der deutschsprachigen Schweiz, in Auftrag der interkantonalen Arbeitsgemeinschaft Externe Schulevaluation (argev) nachgegangen. Ziel des Projekts ist die Entwicklung eines auf den Unterricht nach Lehrplan 21 ausgerichteten, überkantonal einsetzbaren Beurteilungsinstrumentariums zur Erfassung von Unterrichtsqualität auf der Volksschulstufe sowie die Konzeption eines Qualifizierungsangebots für Evaluationsfachpersonen. Aus dem Projekt resultierte das Instrumentarium zur Unterrichtsbeurteilung ausgerichtet auf den Lehrplan 21 im Auftrag von argev (INSULA) mit 7 Dimensionen und 21 Subdimensionen von Unterrichtsqualität (Rogh et al., 2020) mit einer dazugehörigen dreieinhalbtägigen Schulung. Über drei Tage werden theoretische und methodische Grundlagen, die zu beurteilenden (Sub-)Dimensionen von Unterrichtsqualität sowie die Anwendung des Beurteilungsinstrumentariums trainiert. Daraufhin erfolgt eine halbtägige Zertifizierung, bei der Evaluationsfachpersonen mehrere Unterrichtsaufzeichnungen selbstständig beurteilen müssen.

Zwischen 2021-2022 wurden sechs Schulungen für Evaluationsfachpersonen aus 15 Kantonen durchgeführt. Durchschnittlich nahmen 23 Personen (MIN=10; MAX=36) an den Schulungen teil, die pandemiebedingt online durchgeführt wurden. 95.1% der Teilnehmenden waren vor ihrer Evaluationstätigkeit als Lehrperson tätig. Durchschnittlich bringen die Teilnehmenden 14.85 Jahre (MIN=0; MAX=35) Unterrichtserfahrung und 5.13 Jahre (MIN=0; MAX=25) Evaluationserfahrung mit. Mittlerweile liegen die Zertifizierungsergebnisse von 93 Personen vor, die über den Abgleich mit Mastercodierungen entsprechende Anhaltspunkte zur Qualifizierung in der Unterrichtsbeurteilung liefern. Die Übereinstimmungsprüfung zwischen Masterratings und Zertifizierungsratings erfolgt dimensionsspezifisch sowie dimensionsübergreifend. Idealerweise beträgt die exakte Übereinstimmung über alle Ratings mindestens 60% (Bell et al., 2012; Lotz et al., 2013; Wilhelm et al., 2018).

Ergebnisse und ihre Bedeutung

Über alle Schulungen hinweg beträgt die dimensionsübergreifende Übereinstimmung 62% (SD=.55; MIN=50%; MAX=73%). Die dimensionsspezifische Auswertung zeigt, dass Dimensionen wie Klassenführung oder Konsolidieren und Festigen von besonders hohen Übereinstimmungen zeugen (82-90%), wohingegen Dimensionen wie Beurteilen und Feedback oder kognitive Aktivierung besonders herausfordernd scheinen (38%-44%). Nach der dreitägigen Schulung wird also insgesamt lediglich ein Mindestmaß an Übereinstimmung zur Qualifizierung erreicht und die Unterrichtsbeurteilungen bleiben von Verzerrungstendenzen behaftet. Implikationen mit Blick auf die Rolle von Masterratings, aber auch auf die Identifikation von Effekten von Unterrichtsqualität werden diskutiert. Darüber hinaus bleibt offen, inwiefern Entscheidungsfindungsstrategien wirksam trainiert wurden. Überlegungen, wie diesem Aspekt in zukünftigen Schulungen durch begleitende kognitive Interviews begegnet werden kann, werden im Ausblick vorgestellt.

D-12: Internationale Studierende in Deutschland – Interdisziplinäre Befunde zu Studienerleben, Studienerfolg und Verbleibsintentionen

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* T03 R02 D26

Internationale Studierende in Deutschland – Interdisziplinäre Befunde zu Studienerleben, Studienerfolg und Verbleibsintentionen

Chair(s): **Julia Zimmermann** (FernUniversität in Hagen, Germany)

Internationale Studierende, d.h. Studierende ohne deutsche Staatsangehörigkeit, die ihre Hochschulzugangsberechtigung im Ausland erworben haben, leisten einen wichtigen Beitrag zur Internationalisierung des deutschen Hochschulsystems und zur Bewältigung des Fachkräftemangels (Hoffmeyer-Zlotnik & Grote, 2019). Zudem stellen sie mit einem Anteil von 11% an der Studierendenschaft im Studienjahr 2020/21 eine bedeutsame Gruppe an deutschen Hochschulen dar (Statistisches Bundesamt, 2022). Jedoch sind sie im Vergleich zu einheimischen Studierenden seltener erfolgreich in ihren Bemühungen, einen Studienabschluss zu erlangen (Heublein et al., 2022). Trotz einer Intensivierung der Forschungsaktivitäten in den vergangenen Jahren (z.B. Pineda et al., 2022; Wisniewski et al., 2022) liegen im Vergleich zum internationalen Forschungsstand weiterhin vergleichsweise wenige empirische Erkenntnisse zur Studiensituation und den Bedingungsfaktoren des Studienerfolgs dieser Studierenden sowie ihren (beruflichen) Perspektiven in Deutschland vor.

Im Rahmen des Symposiums werden fünf Beiträge präsentiert, die diese Forschungslücke aus dem Blickwinkel unterschiedlicher Disziplinen (Psychologie, Soziologie, Sprach- und Erziehungswissenschaften) adressieren. Dabei werden verschiedene sprachliche, soziale, motivationale und strukturelle Bedingungsfaktoren unterschiedlicher Indikatoren des Studienerfolgs (persönliches Wachstumserleben („Thriving“), Studienzufriedenheit, Studienleistungen) und der Verbleibsintentionen internationaler Studierender betrachtet. Die quantitativ-empirischen Forschungsarbeiten basieren auf unterschiedlichen (Panel)datensätzen und nutzen komplexe Verfahren der (längsschnittlichen) Datenanalyse.

Der erste Beitrag adressiert aus einer intersektionalen Perspektive die Bedeutung von soziodemographischen Merkmalen (Geschlecht, akademischer und kultureller Hintergrund) sowie Akkulturationsorientierungen mit Blick auf die motivationalen Ausgangsbedingungen von internationalen Studierenden zum Beginn ihres Bachelorstudiums in Deutschland in MINT- und nicht-MINT-Studiengängen.

Der zweite Beitrag geht der Frage nach, welche Rolle das Mindset der Lehrenden – neben motivationalen und sozialen Merkmalen der Studierenden – für das persönliche Wachstumserleben („Thriving“) internationaler Studierender in Deutschland während der Corona-Pandemie spielte.

Auch der dritte Beitrag nimmt motivationale Merkmale – hier jedoch Motivationsregulationsstrategien – in den Blick und untersucht die quer- und längsschnittlichen Zusammenhänge zwischen Motivationsregulation, akademischer Anstrengung sowie Studienzufriedenheit und Studienleistungen internationaler Studierender im Verlauf ihres Bachelorstudiums in Deutschland.

Der vierte Beitrag gibt einen Überblick über die zentralen Befunde eines Forschungsprojekts, das sich der Erfassung sprachlicher Studienvoraussetzungen sowie der Bedeutung von sprachlichen Faktoren

im Hinblick auf verschiedene Indikatoren des Studienerfolgs internationaler Studierender im deutschen Hochschulsystem widmete.

Der fünfte Beitrag adressiert aus einer migrationssoziologischen Perspektive die Frage, unter welchen individuellen und strukturellen Bedingungen internationale Studierende nach Abschluss ihres Studiums einen dauerhaften Verbleib in Deutschland anstreben.

In der Zusammenschau bietet das interdisziplinäre Symposium vielfältige, differenzierte und hoch aktuelle Erkenntnisse zur Studiensituation und den Bedingungsfaktoren des Studienerfolgs und der Verbleibsintentionen internationaler Studierender in Deutschland und gibt damit wichtige Hinweise für die (Weiter)entwicklung von Maßnahmen zur Unterstützung dieser Studierendengruppe.

Beiträge des Symposiums

Studienbezogene Selbstwirksamkeit und Wertüberzeugungen internationaler MINT- und Nicht-MINT-Studierender in Deutschland aus intersektionaler Perspektive

Judith Sarah Preuß, Julia Zimmermann, Kathrin Jonkmann
FernUniversität in Hagen

Theoretischer Hintergrund

In Anbetracht des bestehenden Fachkräftemangels ist die Förderung des Studienerfolgs internationaler (MINT-)Studierender in Deutschland von gesamtgesellschaftlichem Interesse (Hoffmeyer-Zlotnik & Grote, 2019). Dazu ist es relevant, die Bedingungsfaktoren des Studienerfolgs dieser Studierendengruppe zu verstehen. Nach der Erwartungs-mal-Wert Theorie (EVT, Eccles et al., 1983) gelten die studienbezogene Selbstwirksamkeit, d.h. die Erwartung der Studierenden, ihr Studienverhalten selbst kontrollieren, reflektieren und die akademischen Anforderungen erfüllen zu können (Bandura et al., 1999), sowie studienbezogene Wertüberzeugungen (Eccles, 1994) wie intrinsic value (Interesse an der Aufgabe), attainment value (persönliche Bedeutung des Erfolgs, Identifikation mit der Aufgabe), utility value (erwarteter Nutzen der Aufgabe für persönliche Ziele) und costs (Aufwand, Opportunitätskosten, emotionale Kosten) als zentrale Voraussetzungen des Studienerfolgs.

Jedoch unterscheiden sich (internationale) Studierende schon zu Studienbeginn in ihren Ausprägungen dieser Merkmale. Vorhergehende Studien zeigten Zusammenhänge zwischen dem Geschlecht (Gaspard et al., 2015; Tellhed et al., 2017), dem akademischen (Blaney & Stout, 2017; Jiang et al., 2020) und dem kulturellen Hintergrund (Litson et al., 2021; Seo et al., 2019) der (MINT-)Studierenden und ihren studienbezogenen Erwartungen und Wertüberzeugungen. In Anbetracht der Herausforderungen der kulturellen Transition erscheinen zudem die Akkulturationsorientierungen internationaler Studierender in dieser Hinsicht relevant.

Fragestellung

Um potentielle Risikogruppen internationaler Studierender mit ungünstigen Ausprägungen studienbezogener Erwartungen sowie Wertüberzeugungen zu identifizieren, analysierte die vorliegende Studie Zusammenhänge zwischen Geschlecht, akademischem Hintergrund, kulturellem Hintergrund und Akkulturationsorientierungen internationaler Studierender in MINT- und Nicht-MINT-Studiengängen sowie deren studienbezogener Selbstwirksamkeit und Wertüberzeugungen zu

Studienbeginn. Aus intersektionaler Perspektive wurden Wechselwirkungen zwischen Geschlecht und akademischem Hintergrund, kulturellem Hintergrund sowie Akkulturationsorientierungen in Bezug auf Unterschiede in den Erwartungs- und Wertkomponenten untersucht.

Methode

Die Daten internationaler Bachelorstudierender (N = 1,544, nMINT = 851, nnicht-MINT = 693) einer deutschen Panelstudie zu internationalen Studierenden (International Student Survey, Falk et al., 2021) wurden für hierarchische Mehrgruppen-Regressionsanalysen (Mplus Version 7; Muthén & Muthén, 1998-2015) genutzt. Mögliche signifikante Unterschiede der Zusammenhänge zwischen MINT- und Nicht-MINT-Studierenden wurden mittels iterativer Wald-Tests geprüft.

Die studienbezogene Selbstwirksamkeitserwartung wurde mit einer adaptierten Version der Kurzskala zur Erfassung allgemeiner Selbstwirksamkeitserwartungen (ASKU, Beierlein et al., 2013) erhoben. Die studienbezogenen Wertüberzeugungen achievement value, utility value und costs wurden mit je zwei Items der Kurzskala von Gaspard et al. (2015) operationalisiert, intrinsic value mit zwei Items der Facette „Zufriedenheit mit den Studieninhalten“ der Studienzufriedenheitsskala (Westermann et al., 1996, in Anlehnung an Heise & Thies, 2015). Zur Erfassung des kulturellen Hintergrundes wurden die Herkunftsländer der Studierenden zu Ländergruppen (DAAD & DZHW, 2020) zusammengefasst. Die Kategorie „mit akademischem Hintergrund“ wurde vergeben, wenn mindestens ein Elternteil einen Hochschulabschluss aufwies.

Ergebnisse und ihre Bedeutung

Lediglich der Zusammenhang zwischen akademischem Hintergrund und attainment value fiel für die Fächergruppen signifikant unterschiedlich aus, d.h. MINT-Studierende mit akademischem Hintergrund und Nicht-MINT-Studierende ohne akademischen Hintergrund zeigten die höchsten Ausprägungen. Alle weiteren Resultate galten für beide Fächergruppen gleichermaßen. Es zeigten sich keine direkten Zusammenhänge zwischen Geschlecht und den Erwartungs- und Wertkomponenten. Für intrinsic value zeigte sich jedoch eine signifikante Interaktion zwischen Geschlecht und akademischem Hintergrund, weibliche Studierende mit akademischem Hintergrund erzielten hier die höchsten Ausprägungen. Des Weiteren wurden signifikante Zusammenhänge zwischen dem kulturellen Hintergrund und allen untersuchten Erwartungs- und Wertkomponenten identifiziert, die meisten Befunde entfielen auf die Variable costs. Einige Analysen detektierten bedeutsame Interaktionen von kulturellem Hintergrund und Geschlecht, was auf Unterschiede zwischen männlichen und weiblichen Studierenden aus einzelnen Ländergruppen hindeutet. Akkulturationsorientierungen korrelierten positiv mit attainment value, utility value und intrinsic value, d.h. eine höhere Gastkulturorientierung ging mit höheren Werten dieser Wertkomponenten einher, während eine höhere Herkunftskulturorientierung lediglich mit höheren intrinsic-Werten zusammenhing.

Die Modelle erklärten substantielle Varianz in allen abhängigen Variablen (7%-13%). Die Ergebnisse sind relevant für die zielgruppenspezifische Umsetzung von Maßnahmen zur Unterstützung studienbezogener Selbstwirksamkeit und adaptiver Wertüberzeugungen und somit letztlich für die Förderung des Studienerfolgs internationaler Studierender in Deutschland.

Thriving von internationalen Studierenden: die Rolle der Mindsets von Dozierenden, sowie der Selbstwirksamkeitserwartungen und des Zugehörigkeitserlebens der Studierenden

Jannika Haase, Lysann Zander
Leibniz Universität Hannover

Theoretischer Hintergrund

Internationale Studierende können den Übergang in ein neues akademisches Umfeld als große Herausforderung erleben. Stress- und Resilienzforschung haben gezeigt, dass Personen auf herausfordernde Situationen mit erhöhtem Stresserleben, Gefühlen von Pessimismus und Hoffnungslosigkeit reagieren (z.B. Lazarus & Folkman, 1984). Im Umgang mit belastenden Situationen zeigen manche Personen jedoch auch adaptive Reaktionen und agieren in ihrem Umfeld auf eine Weise, die nicht nur ihr Wohlbefinden fördern (Resilienz; Leipold & Greve, 2009), sondern sogar von Erfahrungen persönlichen Wachstums begleitet sind (LePine et al., 2004).

Eine Form dieser Wachstumserfahrung ist Thriving, das Erleben von Vitalität, individuellen Lernerfahrungen und persönlicher Entwicklung (z.B. Benson & Scales, 2009; Bundick et al., 2010; Spreitzer et al., 2005) in herausfordernden Lebensumständen. Verschiedene Faktoren können das Erleben von Thriving begünstigen (Brown et al., 2017), wobei zwischen individuellen (insbesondere Selbstwirksamkeit, Niessen et al., 2012) und sozialen (z.B. soziale Beziehungen, Spreitzer et al., 2005) Faktoren differenziert wird.

Während der COVID-19 Pandemie waren soziale Beziehungen in Online-Lernumgebungen besonders schwer zu etablieren. Auch internationale Studierende erlebten negative Gefühle wie akkulturativen Stress oder Einsamkeit (Wilczewski et al., 2021). Als Minoritätsgruppe in ihrer jeweiligen Lernumgebung könnten sie zudem besonders empfänglich für implizit oder explizit formulierte Erwartungen gewesen sein, die von statushöheren Akteur*innen, wie z.B. Dozierenden, vermittelt wurden (Murphy et al., 2007). Diese Erwartungen können beispielsweise die wahrgenommene Einstellung der Dozierenden bezüglich der Veränderbarkeit der Fähigkeiten ihrer Studierenden betreffen. Diese können die Form von Growth Mindset (d.h., Vorstellung, dass Intelligenz durch Anstrengung veränderbar ist) oder eines Fixed Mindsets (d.h., die Vorstellung, dass Intelligenz angeboren und unveränderbar ist) annehmen (Dweck, 1999). Aktuelle Forschungsarbeiten zeigen, dass Growth- bzw. Fixed Mindsets von Lehrkräften akademische Outcomes von Lernenden vorhersagen können (z.B. Heyder et al., 2020). Studierende, die ein Fixed Mindset bei ihren Dozierenden wahrnehmen, können negative Affekte (z.B. Nervosität, Angst, Stress; Muenks et al., 2020) und somit möglicherweise gleichzeitig weniger Vitalität und Wachstum erleben.

Fragestellung

Die vorliegende Studie untersuchte, ob die durch Studierende wahrgenommenen Erwartungen der Dozierenden zum Erleben persönlichen Thrivings beitrugen - über klassische Prädiktoren wie Selbstwirksamkeits-überzeugungen der Studierenden in digitalen Lernumgebungen (individueller Faktor) und die soziale Zugehörigkeit (sozialer Faktor) hinaus.

Methode

Die Analysen basierten auf den Angaben von 131 internationalen Studierenden im Frühjahr 2021, die zu Beginn des zweiten Online-Semesters an der Befragung einer großen Universität in Deutschland

teilnahmen. 76 Teilnehmende wären männlich, 49 weiblich und eine Person divers (n=5 keine Angabe). Das durchschnittliche Alter der Teilnehmenden lag bei 25.34 (SD=3.55); 56 Personen studierten im Master und 26 im Bachelor (n=49 keine Angabe).

Das Messinstrument für Thriving (Porath et al., 2012) beinhaltet eine Vitalitätsdimension (z.B. „Ich fühle mich lebendig und vital.") und eine Lerndimension (z.B. „Ich lerne gerade sehr viel dazu."). Zudem wurden akademische Online-Selbstwirksamkeit (Shen et al., 2013), die wahrgenommene Zugehörigkeit zum Studiengang (Zander & Höhne, 2021) sowie die durch Studierende wahrgenommenen Growth Mindsets der Dozierenden (Spinath & Schöne, 2003; Dweck, 1999) erfasst.

Ergebnisse und ihre Bedeutung

Eine hierarchische Regressionsanalyse mit full information maximum likelihood (FIML)-Schätzung für fehlende Werte wurde in Mplus Version 8.5 (Muthén & Muthén, 1998–2017) gerechnet. Es zeigte sich, dass die Vorstellung der Dozierenden, dass Intelligenz veränderbar ist, das Thriving der Studierenden positiv vorhersagte (β=.141, p=≤.01), und dies über die akademische Online-Selbstwirksamkeit (β=.422, p=≤.001), und die wahrgenommene Zugehörigkeit zum Studiengang hinaus (β=.254, p=≤.01).

Auch wenn diese ersten Ergebnisse der Replikation bedürfen, legen sie doch nahe, dass Dozierenden eine bedeutsame Rolle dabei zukommen kann, internationale Studierende im Umgang mit den Herausforderungen des Übergangs in ein anderes Bildungssystem zu unterstützen. Es ist davon auszugehen, dass diese Herausforderungen durch die Online-Lehre während der Pandemie verstärkt wurden. Wir diskutieren die Ergebnisse im Vergleich zu Ergebnissen in nicht-internationalen Studierendensamples und diskutieren Implikationen für die Gestaltung von Lernumgebungen.

Motivationsregulationsstrategien, akademische Anstrengung und Studienerfolg internationaler Studierender in Deutschland

Hüseyin Hilmi Yildirim, Julia Zimmermann, Kathrin Jonkmann
FernUniversität in Hagen

Theoretischer Hintergrund

Weltweit beginnen immer mehr Studierende ein Hochschulstudium außerhalb ihres Herkunftslandes. Im Jahr 2019 betraf dies 6,1 Millionen Lernende (OECD, 2021). Internationale Studierende, also ausländische Studierende, die ihre Hochschulzugangsberechtigung im Ausland oder an einem Studienkolleg erworben haben (Statistisches Bundesamt, 2022), sind auch eine wichtige Studierendengruppe in Deutschland. Obwohl die Mehrheit einen Studienabschluss in Deutschland plant, weisen sie im Vergleich zu deutschen Studierenden besonders hohe Abbruchquoten auf (Heublein et al., 2022). Aufgrund des demografischen Wandels, des Fachkräftemangels und wirtschaftlicher Interessen ist die Untersuchung der Faktoren, die den Studienerfolg internationaler Studierender in Deutschland bestimmen, besonders relevant (Morris-Lange, 2019).

Bisherige Studien bestätigten den positiven Zusammenhang zwischen selbstreguliertem Lernen und akademischen Leistungen (Seli & Dembo, 2020). In selbstregulierten Lernmodellen wird zwischen kognitiven, metakognitiven und motivationalen Selbstregulationsstrategien unterschieden (z.B. Boekaerts, 1996). Dabei liegt der Fokus motivationaler Regulationsstrategien auf Gedanken,

Handlungen oder Verhaltensweisen, die dazu dienen, mehr Energie in die anstehende Aufgabe zu investieren und die Beharrlichkeit bei der Erledigung von Aufgaben zu steigern (Schwinger et al., 2007). Die Motivationsregulationsstrategien können somit als eine Form des Ressourcenmanagements verstanden werden. Bisherige (überwiegend querschnittliche) Befunde legen nahe, dass die Nutzung von Motivationsregulationsstrategien im Zusammenhang mit besseren Noten und höherer Studienzufriedenheit bei einheimischen Studierenden steht (Heinze, 2008) und dass diese Korrelationen teilweise mit einer erhöhten akademischen Anstrengung zu erklären sind (z.B. Kryshko et al., 2020).

Fragestellung

Die vorliegende Studie untersuchte die direkten und indirekten Zusammenhänge zwischen Motivationsregulationsstrategien, akademischer Anstrengung und Studienerfolg (Notendurchschnitt und Studienzufriedenheit) internationaler Studierender in Deutschland sowohl im Querschnitt als auch im Längsschnitt. Dabei wurden die inkrementellen Effekte der Motivationsregulationsstrategien über demographische und studienbezogene Merkmale wie Alter, Geschlecht, Abiturnote, vorheriger Deutschlandaufenthalt, Herkunftsregion, Studiensprache und studienbezogene Sprachkenntnisse hinaus betrachtet.

Methode

Stichprobe und Durchführung. Die Daten stammten aus einer Panelstudie (t1–t6; 3 Jahre) zu internationalen Studierenden in Deutschland (International Student Survey, Falk et al., 2021). Die Analysen umfassten 1,060 internationale Bachelorstudierende aus 101 Ländern, die auch zu t5 noch in Deutschland eingeschrieben waren. Die Teilnehmenden studierten an 110 Hochschulen in 34 Fächergruppen. Das mittlere Alter zu t1 betrug 25.35 Jahre und 52.8% der Stichprobe war weiblich. Die Auswertung erfolgte anhand von Mplus (Version 7; Muthén & Muthén, 1998-2015) mittels drei verschiedener Analysemodelle: 1. querschnittliche Mediationsmodelle, 2. längsschnittliche Mediationsmodelle ohne Kontrolle autoregressiver Pfade und 3. längsschnittliche Mediationsmodelle mit Kontrolle autoregressiver Pfade. Um die indirekten Effekte zu testen, wurden Bias-korrigierte Bootstrap-Konfidenzintervalle basierend auf 1000 Stichproben verwendet. Es wurden separate Modelle für den Gesamtmittwert der Motivationsregulationsstrategien sowie alle einzelnen Strategien berechnet.

Messinstrumente. Die Motivationsregulationsstrategien und die akademische Anstrengung wurden mit adaptierten Items aus den Inventaren von Schwinger et al. (2007) und Boerner et al. (2005) erfasst. Die Studienzufriedenheit wurde mit einer angepassten Version der Skala von Westermann et al. (1996) erhoben, während die Studiendurchschnittsnote mit einem Item aus dem NEPS erfasst wurde.

Ergebnisse und ihre Bedeutung

Der Gesamtmittelwert der Motivationsregulationsstrategien sowie die Einzelstrategien Umweltkontrolle und Zielsetzung zeigten in quer- und längsschnittlichen Modellen ohne Kontrolle autoregressiver Pfade über die akademische Anstrengung vermittelte Zusammenhänge mit der Studiendurchschnittsnote und der Studienzufriedenheit. Die Strategien Steigerung der persönlichen Bedeutsamkeit und Selbstbelohnung wiesen nur im Querschnitt indirekte Zusammenhänge mit beiden

Studienerfolgskriterien auf. Die Strategien leistungszielbezogene und vermeidungs-leistungszielbezogene Selbstinstruktionen zeigten ebenfalls nur im Querschnitt indirekte Zusammenhänge mit der Studienzufriedenheit. Die Strategie leistungszielbezogene Selbstinstruktion hatte im Längsschnitt (ohne Kontrolle autoregressiver Pfade) und die Strategie vermeidungs-leistungszielbezogene Selbstinstruktion im Querschnitt einen indirekten Zusammenhang mit der Studiendurchschnittsnote. In den längsschnittlichen Modellen mit Kontrolle der autoregressiven Pfade wurden hingegen weder für den Gesamtmittelwert der Strategienutzung noch für eine der einzelnen Strategien bedeutsame indirekte Zusammenhänge mit den Studienerfolgskriterien ermittelt. Implikationen der Befunde für das wissenschaftliche Verständnis der Bedeutung von Motivationsregulationsstrategien für den Studienerfolg (internationaler) Studierender werden abschließend diskutiert.

Deutschkompetenzen: Ein Schlüssel zum Studienerfolg internationaler BA-Studierender

Katrin Wisniewski
Otto-Friedrich-Universität Bamberg

Theoretischer Hintergrund

Vor dem pandemiebedingten Rückgang hatte die Zahl der internationalen Studierender an deutschen Hochschulen stetig zugenommen. Allerdings brechen diese Studierenden sehr häufig ihr Studium ab, v.a. in Bachelorstudiengängen (vgl. Heublein et al., 2020: 49%). Bis zur Durchführung der Projekte SeSaBa (Pineda et al., 2022) und des hier fokussierten Projekts Sprache und Studienerfolg bei Bildungsausländer/-innen (SpraStu) lagen nahezu keine empirischen Studien zu möglichen Ursachen vor.

SpraStu nahm sprachliche Faktoren für den geringeren Studienerfolg von Bildungsausländer:innen in deutschsprachigen Studiengängen in den Blick. Hintergrund war die weit verbreitete, empirisch jedoch lückenhaft (nur deskriptiv) unterfütterte Vermutung, dass der mangelnde Studienerfolg dieser Gruppe neben akkulturationsbedingten und finanziellen v.a. sprachliche Ursachen haben könnte (z.B. Sozialerhebungen, siehe auch DAAD/DZHW, 2019, S. 156–159; Li, 2017; Zhang et al., 2010; vgl. Morris-Lange, 2017 und Wisniewski et al. 2022, S. 31-36). Erst in jüngster Zeit haben Publikationen aus SeSaBa auch ursächliche Faktoren des Studienerfolgs von Bildungsausländer:innen in Deutschland aufgezeigt (z.B. Pineda et al., 2019; Thies & Falk, 2021; Yildirim et al., 2021a,b), dabei allerdings keine sprachlichen Faktoren analysiert. Internationale Studien, deren Ergebnisse sich aus verschiedenen Gründen nicht einfach auf die deutsche Situation übertragen lassen, finden meist geringe (teils auch keine) Effekte von Sprachkompetenzen auf Studienerfolg. Aussagekräftigere Ergebnisse ergeben sich bei Nutzung fertigkeits- bzw. sprachbereichsspezifischer Testverfahren und differenzierter Betrachtung der Zielgruppe (vgl. Bridgeman et al., 2015; Cho & Bridgeman, 2012; Harsch et al., 2017; Macdonald, 2019; vgl. Gagen, 2019; Pearson, 2021; Wisniewski et al., 2022, S. 48-55).

Fragestellung

Der Vortrag fasst zunächst die o.g. Ausgangssituation zusammen und geht auf sprachliche Bedingungen des Studiums internationaler Studierender ein. Dann werden Hauptergebnisse aus SprasStu zu folgenden Fragen vorgestellt:

1. Wie sprachlich kompetent starten Bildungsausländer:innen ins Studium und wie entwickeln sich ihre Sprachkompetenzen?

2. Sagen Sprachkompetenzen Studienerfolg vorher?

3. Wie ist das relative prädiktive Gewicht von Deutschkompetenzen im Vergleich zu anderen Studienerfolgsprädiktoren?

Methoden

Für N = 340 L2-Bildungsausländer:innen, immatrikuliert an den Projektuniversitäten Leipzig und Würzburg in deutschsprachigen BA- oder Staatsexamensstudiengängen im ersten Fachsemester, wurde in einem kreuzversetzen longitudinalen Paneldesign ein Mixed-Methods-Ansatz verfolgt. Befunde der hier zu präsentierenden Analysen beruhen auf den Ergebnissen a) eines Vorbedingungsfragebogens, b) von sechs standardisierten, jährlich eingesetzten Sprachtests, und c) einem halbjährlich ausgeteilten Studienbedingungsfragebogen. Es wurde eine Reihe subjektiver (erfragte) und objektiver (über die Studienverwaltung vermittelte) Studienerfolgsindikatoren untersucht. Zur Analyse wurden deskriptive Verfahren und je nach statistischen Voraussetzungen schrittweise hierarchische Regressionen bzw. gemischt-lineare Modelle bzw. ein Pfadmodell gerechnet.

Ergebnisse und ihre Bedeutung

FF1: Trotz identischer Studieneingangsbedingungen (Sprachtests auf Niveau B2.2/C1.1) verfügten die Bildungsausländer:innen zum Studienstart über heterogene Sprachkompetenzen. Knapp 40% erreichten teils nicht das Niveau B2, 10% nie. Es zeigten sich Zusammenhänge mit der Herkunftsregion, der Sprachlernbiografie, der Wirtschaftskraft des Herkunftslands und der studierten Fächergruppe. Außerdem unterschied sich die Sprachkompetenz je nach Hochschulzugangssprachtest. Die Sprachkompetenzen verbesserten sich meist im Studienverlauf. Nach zwei Jahren erreichten 50–60% der Bildungsausländer:innen C1. Wer sprachlich gut gerüstet ins Studium einstieg, behielt im Studienverlauf sprachliche Vorsprünge.

FF2: Sprachkompetenzen zu Studienbeginn, v.a. das Leseverstehen, waren prädiktiv für den objektiven Studienerfolg im ersten Studienjahr. Weniger kompetente L2-Studierende waren häufiger unzufrieden mit ihrem Studium und fühlten sich eher überlastet. Später im Studium ließ sich nur noch ein Einfluss auf den objektiven Studienerfolg finden (über den Beobachtungszeitraum je ca. 20%). Andere sprachliche Konstrukte (language mindsets, willingness to communicate, sprachliche Kontakte auf Deutsch) bedingten wiederum subjektiven Studienerfolg.

FF3: In einer hierarchischen Regressionsanalyse und einem Strukturgleichungsmodell erwies sich, dass sich ein gutes Drittel (36.5%) der Varianz in den Leistungen im ersten Studienjahr durch Lesekompetenzen, die soziale und akademische Integration sowie durch das Gefühl, mit den eigenen finanziellen Mitteln auszukommen, erklären ließen. Andere, etwa motivationale oder auf das akademische Selbstverständnis bezogene Faktoren oder aber auch die generelle Problembelastung hingegen spielten keine statistisch signifikante Rolle.

Integration, berufliche Perspektiven und Verbleibsintentionen internationaler Studierender in Deutschland

Knut Petzold, Marie Siebenhaar
Hochschule Zittau/ Görlitz

Theoretischer Hintergrund

Vor dem Hintergrund des demografischen Wandels und eines daraus hervorgehenden Fachkräftemangels zeigen sich zahlreiche Herausforderungen für Wirtschaft und Sozialstaat in Deutschland (Destatis, 2019). Ohne Fachkräfte aus dem Ausland sind diese zukünftig nicht zu bewältigen (BMAS, 2021). Ein besonderes Potenzial bieten dabei internationale Studierende, wenn diese im Anschluss an ihren Studienabschluss erfolgreich in den deutschen Arbeitsmarkt integriert werden können (DAAD & DZHW, 2021). Diese haben durch ihr vorheriges Studium häufig bereits sprachliche, kulturelle und soziale Kompetenzen sowie arbeitsmarktspezifische Erfahrungen gemacht und befinden sich bereits in Deutschland, wodurch bürokratische Prozesse vereinfacht und Integration sowie Akkulturation bereits fortgeschritten sind (Morris-Lange, 2019).

Unter Rückgriff auf das in der Hochschulforschung etablierte Humankapitalmodell der Migration (z.B. Sjasstad, 1962; DaVanzo, 1981) kann der Verbleib als fortgesetzte Migrationsentscheidung konzeptualisiert werden, der eine erneute Investition in die eigene Zukunft darstellt. Demnach müssen im Rahmen eines subjektiven Kosten-Nutzen-Kalküls die erwarteten Erträge akademischer Bildung beim Erreichen der eigenen Ziele die wahrgenommenen Kosten eines Verbleibs in Deutschland übersteigen. Während die wahrgenommenen Erträge vor allem durch die beruflichen Entwicklungsmöglichkeiten repräsentiert werden, korrespondieren die Kosten des Verbleibs insbesondere mit den Aspekten der strukturellen, kulturellen, sozialen und emotionalen Integration in Deutschland und im Herkunftsland (Heath & Schneider, 2021).

Bisher liegen Studien vor allem für den angloamerikanischen Hochschul- und Arbeitsmarktkontext vor (Bijwaard & Wang, 2016; Cheung & Xu, 2015; Feinberg, 2021; Soon, 2010; Yu, 2016). Die Ergebnisse deutscher Studien sind durch die ausschließliche Nutzung von Registerdaten beschränkt (Geis, 2017) oder weil bestimmte Standorte (Hooijen et al, 2017; Koenings et al., 2020), Hochschulformen bzw. Studiengänge (Sykes & Chaoimh, 2012) fokussiert werden.

Fragestellung

In diesem Beitrag wird daher die Frage verfolgt, unter welchen individuellen und strukturellen Bedingungen internationale Studierende nach Abschluss ihres Studiums einen dauerhaften Verbleib in Deutschland anstreben (Petzold, 2022).

Methode

Die theoretischen Überlegungen wurden mit Daten aus groß angelegten Studierendenbefragungen überprüft, die aufgrund der regionalen und institutionellen Breite umfassende Rückschlüsse auf internationale Studierende in Deutschland insgesamt zulassen, aber auch vertiefte Subgruppenanalysen (nach Fächergruppen, Hochschultyp, Region und Herkunftsland) erlauben.

Die Datenanalyse wurde mithilfe fortgeschrittener Regressionsverfahren durchgeführt (Best & Wolf, 2015). Als Prädiktoren des Verbleibs wurden Variablen hinzugezogen, welche erwartete Erträge, subjektive Sicherheit, strukturelle Bedingungen, kulturelle sowie soziale Ressourcen und emotionale

Bewertungen abbilden. Diese wurden stufenweise einbezogen, um deren relative Bedeutung für die Verbleibsintention abzuschätzen sowie um zu analysieren, inwiefern Effekte erwarteter Erträge durch strukturelle, kulturelle, soziale und emotionale Ressourcen vermittelt werden. Die Analyse von Interaktionstermen wurde darüber hinaus genutzt, um die Heterogenität der Effekte erwarteter beruflicher Erträge zu betrachten.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass eine deutliche Mehrheit der internationalen Studierenden eine klare Intention zeigt, nach Studienabschluss in Deutschland zu verbleiben. Dennoch ist ein erkennbarer Anteil in der Frage des Verbleibs noch unentschieden. Ein zentrales Ergebnis ist, dass gute berufliche Perspektiven als erwartete Erträge des Aufenthalts starke positive und über die Modelle hinweg stabile Effekte auf die Verbleibsabsicht aufweisen. Auch die emotionale Bewertung, wie die subjektive Zufriedenheit mit Deutschland als Gastland, zeigt einen positiven Zusammenhang mit der Absicht in Deutschland zu bleiben. Weitere positive Effekte auf die Verbleibsintention zeigen sich auch in Bezug auf soziale Unterstützung im Studium sowie der Zufriedenheit mit der Hochschule. Studierende, welche mit ihrem Profil der fachlichen Nachfrage des Arbeitsmarkts entsprechen bzw. praxisnah studiert haben, neigen darüber hinaus eher zum Verbleib. Es liegen darüber hinaus Unterschiede über institutionelle Kontexte, wie Hochschultyp und Aufenthaltsdauer, vor.

Insgesamt bestätigen die Analysen die große Bedeutung von wahrgenommenen Chancen für einen erfolgreichen Eintritt in den Arbeitsmarkt, die durch kontextuelle Rahmenbedingungen gefördert oder gehemmt werden können. Insbesondere die Bedeutung der sozialen und kulturellen Einbettung im Gastland sollte künftig jedoch differenzierter betrachtet werden. Auch der Bedeutung regionaler Bedingungen im Herkunfts- wie auch Gastland sollte weiter nachgegangen werden.

D-13: Situationale Emotionen und Motivation von Studierenden

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* T03 R02 D39

Situationale Emotionen und Motivation von Studierenden

Chair(s): **Lina Wirth** (Leuphana Universität Lüneburg, Deutschland), **Poldi Kuhl** (Leuphana Universität Lüneburg), **Timo Ehmke** (Leuphana Universität Lüneburg)
*Diskutant*in(nen):* **Lena Wieland** (DIPF | Leibniz Institute for Research and Information in Education)

Emotionen und Motivation gelten als wesentliche Prädiktoren individueller Lernerfolge, da sie in vielfältiger Weise auf das Lernen einwirken (Götz, Frenzel & Pekrun, 2009). Zur Erklärung emotionaler und motivationaler Prozesse in Lernsituationen haben sich insbesondere die Control-Value-Theory (Pekrun, 2006) und die (Situated) Expectancy-Value-Theory (Eccles et al., 1983; Eccles & Wigfield, 2020) etabliert. Die Theorien postulieren, dass lern- und leistungsbezogene Emotionen und Motivation durch Kontrollüberzeugungen und Erfolgserwartungen der Lernenden sowie durch den subjektiven Wert, die diese einem Lerngegenstand zuschreiben, determiniert werden. Die Kontrollüberzeugungen und Erfolgserwartungen sowie Wertzuschreibungen entwickeln sich wiederum in Abhängigkeit individueller sowie umweltbezogener und situationaler Merkmale des Lerngegenstandes (Pekrun, 2006; Eccles et al., 1983).

Während in der Forschung Emotionen und Motivation in diesem Kontext bislang vorwiegend als eigenschaftsähnliche Konstrukte sowie allgemeinere Konzepte der Person-Umwelt-Passung untersucht wurden, bewegt sich der Fokus aktueller Forschung zunehmend in Richtung der Untersuchung dynamischer und situationsabhängiger emotionaler und motivationaler Prozesse (Grassinger et al. 2019; Pekrun & Marsh, 2022; Krannich et al., 2022), zu denen jedoch nach wie vor weitreichender Forschungsbedarf besteht (Eccles, 2022).

Diese aktuellen Entwicklungen und Forschungsdesiderate bei der Untersuchung von situationsbezogenen Emotionen und Motivation Lernender adressierend, hat dieses Symposium zum Ziel, Forschungsarbeiten zu Emotionen und Motivation im Kontext situationaler Merkmale des Lernens zusammenzubringen.

Vor diesem Hintergrund beschäftigt sich der erste Beitrag des Symposiums mit der situationalen Variabilität in der inhaltlichen Klarheit von Lehrpersonen und untersucht, wie diese mit der Erwartungs-Wert-Motivation von Studierenden zusammenhängt.

Der zweite Beitrag beleuchtet die Lehrdimensionen Lehrerenthusiasmus, Lehrerhumor, Eintönigkeit und Klarheit und untersucht, wie sich die Wahrnehmung dieser Lehrdimensionen auf das Freudeerleben von Studierenden auswirkt. Ziel ist die Aufschlüsselung stabiler und variabler Komponenten der erlebten Freude sowie der perzipierten Lehrdimensionen. Zudem werden die dynamische Natur und interaktive Prozesse erlebter Freude sowie von den Studierenden perzipierten Lehrdimensionen beleuchtet und reziproke Beziehungen zwischen Lehrdimensionen und erlebter Freude untersucht.

Der dritte Beitrag des Symposiums untersucht schließlich, wie sich die situationale Langeweile von Studierenden über eine ausgewählte Lehr-Lernsituation hinweg entwickelt und inwieweit sich dabei heterogene Entwicklungsverläufe der Langeweile identifizieren lassen. In einem weiteren Schritt untersucht der Beitrag, ob die Passung sprachlicher Anforderungen und der individuellen sprachlichen Kompetenzen mit der Entwicklung der situationalen Langeweile zusammenhängt.

Die Beiträge sollen abschließend hinsichtlich der Förderung optimaler emotionaler und motivationaler Bedingungen für Studierende diskutiert werden.

Beiträge des Symposiums

Motivation von Studierenden und inhaltliche Klarheit der Lehrenden: Zusammenhänge von Moment zu Moment

Alina Oschwald[1], Bärbel Kracke[1], Julia Moeller[2], Jaana Viljaranta[3], Julia Dietrich[1]
[1]Friedrich-Schiller-Universität Jena, [2]Universität Leipzig, [3]University of Eastern Finland

Theoretischer Hintergrund

Die Motivation von Lernenden variiert erheblich von Moment zu Moment und wird von Lernsituationsmerkmalen mitbestimmt (Martin et al., 2015). Aus Sicht der Erwartungs-Wert-Theorie der Leistungsmotivation (Eccles & Wigfield, 2020) sind besonders die aufgabenspezifischen Erwartungen (Kann ich das?) und subjektiven Werte von Lernenden relevant, wobei die Wertkomponente (Will ich das?) sich in positive Werte wie Nützlichkeits- und intrinsischen Wert und negative Kostenwerte unterteilen lässt.

Auch in Bezug auf Unterrichtsqualität und das Verhalten von Lehrpersonen hat sich der Blick zuletzt auf deren situationalen Variabilität gerichtet (z.B. Voss et al., 2022). Die vorliegende Studie fokussiert auf die inhaltliche Klarheit der Lehrperson. Ein Unterricht mit hoher Klarheit zeichnet sich dadurch aus, dass die Lehrperson Inhalte verständlich, akkurat und kohärent darstellt (Stronge et al., 2011). Eine hohe instruktionale Klarheit hilft Lernenden beim Aufbau von Wissensstrukturen und steht im Zusammenhang mit ihrer Motivation (Brophy, 2000; Lazarides et al., 2019; Maulana et al., 2016).

In dieser Studie stehen die inhaltlichen Aspekte von Klarheit im Vordergrund (Lotz et al., 2013), die wir mit drei Facetten betrachten: 1. dem Detailreichtum der Erklärung, 2. dem Variantenreichtum der Erklärung und 3. der logischen Konsistenz der Argumentation.

Fragestellung

Wir untersuchen im Kontext einer Vorlesung Zusammenhänge zwischen der Erwartungs-Wert-Motivation von Studierenden und der inhaltlichen Klarheit der Lehrperson. Bezogen auf konkrete Lernsituationen vermuteten wir, dass Studierende höhere Erwartungen und Werte sowie geringere Kosten in solchen Lernsituationen angeben, in denen die Lehrperson Inhalte kohärent darstellt und diese in einer variantenreichen, detaillierten Weise präsentiert.

Wir nahmen an, dass inhaltliche Klarheit dafür sorgt, dass Studierende die Inhalte besser verstehen, sie als weniger schwierig wahrnehmen (Eccles & Wigfield, 2020) und sich damit beim Lernen kompetenter fühlen (Deci & Ryan, 1985). Dies sollte sich positiv auf Erwartungen und den intrinsischen Wert auswirken. Auch das Konstruieren von Relevanz der Lerninhalte (Nützlichkeit) könnte bei hoher Klarheit leichter fallen. Demgegenüber könnte eine geringe Klarheit die kognitive Belastung des Arbeitsgedächtnisses erhöhen und die wahrgenommenen Kosten steigern (Feldon et al., 2019).

Methode

Die Studierenden (N = 155, 51% weiblich, Alter: M = 21.77 Jahre, SD = 2.91) wurden in einem short-term intensive longitudinal Design mehrmals pro Vorlesungssitzung über neun Wochen hinweg per

Experience Sampling Methode zu ihrer Motivation befragt (1 = stimmt gar nicht, 4 = stimmt genau; Erwartung: 2 Facetten/Items, z.B. „Ich werde diese Inhalte gut in der Klausur können“; positiver Wert: 3 Items; negativer Wert: 3 Items). Zeitgleich wurden Videoaufzeichnungen der Lehrperson angefertigt und in 3-Minuten-Intervallen durch zwei Rater hochinferent kodiert (0 = im Intervall nicht beobachtet, 1 = teilweise beobachtet, 2 = kontinuierlich beobachtet).

Beide Datensätze wurden so zusammengebracht, dass die Studierendenantworten zu einem bestimmten Signal mit den Videoratings zum gleichen Zeitpunkt korrespondierten.

Die Daten wurden mittels Cross-Classified Multilevel Analysen ausgewertet, bei denen die Beobachtungen des motivationalen Zustands (Level 1, N = 2227) in die Studierenden (Level 2a, N = 155) und den Zeitpunkt der Messung (Level 2b, N = 81) genestet waren. Auf dem Level des Messzeitpunktes (2b) wurde inhaltliche Klarheit in die Modelle aufgenommen. Es wurden Hauptanalysen für die Komponenten Erwartung, positiver Wert, negativer Wert durchgeführt und Zusatzanalysen separat für jede der acht Einzelfacetten.

Ergebnisse und Bedeutung

Die Ergebnisse zeigten keine Zusammenhänge zwischen (objektiv beobachteter) inhaltlicher Klarheit und den übergreifenden Erwartungs-Wert-Komponenten. Wurden Einzelfacetten der Motivation fokussiert, war der Detailreichtum der Lehrperson mit höherer Erfolgserwartung der Studierenden in Bezug auf ihre Klausurleistungen assoziiert, und mit geringeren Anstrengungskosten. Die Befunde stehen im Kontrast zu einigen Studien, die die Wahrnehmung von Instruktionsqualität durch die Lernenden untersuchten (Lazarides et al., 2019; Maulana et al., 2016). Weitere hier nicht untersuchte Bedingungen und Mechanismen, die hinter den Befunden stehen könnten, werden im Beitrag diskutiert.

Beziehungen verschiedener Lehrdimensionen mit dem Freudeerleben in Vorlesungen – ein latenter trait-state-Ansatz

Sonja Bieg[1], Ulrike E. Nett[2], Markus Dresel[2], Thomas Götz[3]
[1]Pädagogische Hochschule Weingarten, [2]Universität Augsburg, [3]Universität Wien

Theoretischer Hintergrund

Im Kontroll-Wert-Ansatz wird postuliert, dass Leistungsemotionen im Unterricht durch individuelle, umweltliche und situationale Faktoren ausgelöst werden. (Pekrun, 2006, 2017). Derzeit ist ein gesteigertes Interesse am Einfluss der Lernumgebung auf die Leistungsemotionen von Lernenden wahrzunehmen, vor allem an der schülerperzipierten Wahrnehmung von Lehrdimensionen (z.B. M. Bieg et al., 2017; S. Bieg et al., 2019; Goetz et al., 2020; S. Bieg et al., 2022). Die generelle Annahme, dass Lehren im Zusammenhang mit den erlebten Schüleremotionen steht, ist empirisch belegt (z.B. Gogol et al., 2017). Wie sich zentrale Lehrdimensionen wie wahrgenommener Lehrerenthusiasmus, Lehrerhumor, Eintönigkeit und Klarheit auf das Freuderleben auswirken, ist noch nicht umfassend untersucht.

Soweit wir wissen, fehlt es an quantitativer Evidenz zu reziproken Beziehungen mehrerer schülerperzipierter Lehrdimensionen und der erlebten Freude bei Studierenden in kurzen reziproken Zyklen, welche die interaktiven Beziehungen dieser Konstrukte in mehreren aufeinander folgenden Vorlesungen untersucht. Zusätzlich berücksichtigen wir, dass Freude gleichzeitig von

personspezifischen (trait) und situationsspezifischen (state) Komponenten beeinflusst wird (z.B. Nett, M. Bieg, & Keller, 2017).

Fragestellung

Ziele der Studie sind (1) stabile und variable Komponenten der erlebten Freude Studierender und deren perzipiertem Enthusiasmus, Humor, Eintönigkeit und Klarheit in Vorlesungen aufzuschlüsseln, (2) die dynamische Natur und interaktive Prozesse der erlebten Freude sowie von den Studierenden perzipierten genannten Lehrdimensionen zu beleuchten und reziproke Beziehungen zwischen diesen Lehrdimensionen und erlebter Freude zu untersuchen.

Methode

An dieser Studie nahmen 559 Studierende aus drei Universitäten in Süddeutschland teil. Die Mehrheit studierte Lehramt (69%), war weiblich (76%) und durchschnittlich 21.6 Jahre alt (SD = 4.49). Die Datenerhebung fand über einen Zeitraum von 4 Wochen in sieben verschiedenen Vorlesungen bei sieben verschiedenen Dozierenden statt mit je einer Vorlesung pro Woche. Innerhalb der Vorlesung gab es drei randomisierte Erhebungszeitpunkte wonach es in Summe pro Studierende/n 12 situative Messungen gab. Um die Emotion Freude und die wahrgenommenen Lehrdimensionen zu messen, wurden Einzelitems genutzt, die von 1 (überhaupt nicht; stimmt gar nicht) bis 5 (sehr stark; stimmt genau) bewertet wurden.

Entsprechend den Zielen, schätzten wir univariate (Ziel 1) und bivariate (Ziel 2) STARTS Modelle (Kenny & Zautra, 2001; Kenny & Zautra, 1995). Alle Modelle wurden mit Mplus 8 analysiert. Die Modelle wurden entsprechend der empfohlenen Fit-indizes von Hu and Bentler (1999) interpretiert.

Ergebnisse und ihre Bedeutung

Bei Freude gehen 30.5% der Gesamtvarianz auf die stabile trait Komponente zurück (person-spezifisch); 31.7% der Gesamtvarianz liegt in den früheren Erfahrungen, die sich teilweise auf frühere states beziehen; 37.8% der Gesamtvarianz beinhaltet einen variablen state Anteil (situations-spezifisch). Wir fanden starke Übertragungseffekte bei Freude von einer Situation auf die nächste und von einer Vorlesung zur nächsten. Bezüglich dem perzipierten Lehrerenthusiasmus und Lehrerhumor entfallen 40.6% bzw. 40.7% der Gesamtvarianz auf einen stabilen trait; 40.6% bzw. 28% auf frühere Erfahrungen, 18.7% bzw. 31.3% waren situationsspezifisch. Enthusiasmus, Eintönigkeit und Klarheit zeigten Übertragungseffekte innerhalb der Vorlesungen, während Humor Übertragungseffekte zusätzlich zwischen den Vorlesungen zeigte. Auf die Eintönigkeit und Klarheit entfallen 30.5% bzw. 48.3% der Gesamtvarianz auf die stabile trait Komponente; 31.5% bzw. 14 % gehen auf vorherige Erfahrungen zurück und 37.9% bzw. 37.8% der Gesamtvarianz sind situationsspezifisch.

In Bezug auf das zweite Ziel zeigte sich kein cross-lagged-Effekt des Erlebens von Freude zu einem spezifischen Zeitpunkt auf die Wahrnehmung der Lehrdimensionen zu einem späteren Zeitpunkt. Jedoch ergab sich ein cross-lagged-Effekt bei Lehrerhumor und Eintönigkeit auf die Freude zu einem späteren Zeitpunkt innerhalb derselben Vorlesung. Lehrerhumor und Klarheit einer Vorlesung zeigten negative cross-lagged-Effekte auf das Freudeerleben in der nächsten Vorlesung, die Eintönigkeit einen positiven cross-lagged-Effekt auf das Freudeerleben in der nächsten Vorlesung.

Scheinbar ist Lehrerhumor ein interaktiveres Element als Lehrerenthusiasmus und Klarheit der Instruktion scheint Erwartungen zu steigern, die folglich nur enttäuscht werden können. Für die Eintönigkeit gilt das Gegenteil.

Situationale Langeweile von Studierenden beim Betrachten eines Erklärvideos: Typologie und Effekte des Passungsverhältnisses von sprachlichen Anforderungen und Kompetenzen

Lina Wirth[1], Poldi Kuhl[1], Timo Ehmke[1], Jan Retelsdorf[2], Burak Aydin[1]
[1]Leuphana Universität Lüneburg, [2]Universität Hamburg

Theoretischer Hintergrund

Langeweile ist im Lern- und Leistungskontext eine häufig erlebte Emotion, die sich negativ auf die Aufmerksamkeit, motivationale Merkmale und damit auf die Leistungen von Lernenden auswirkt (Pekrun et al., 2010). Als situationale Emotion entwickelt sich Langeweile primär durch Über- oder Unterforderung in Lern- und Leistungssituationen. So besagt die Control-Value-Theory von Pekrun (2006), dass Langeweile einerseits vom Gefühl der Kontrolle über lern- und leistungsbezogene Aktivitäten und Ergebnisse und andererseits vom subjektiven Wert bedingt wird, der diesen Aktivitäten und Ergebnissen zugeschrieben wird. Dabei stellen aufgabenbezogene Anforderungen der Lernumgebung eine wesentliche Determinante dieser Kontrollüberzeugungen und Wertzuschreibungen dar (Pekrun et al., 2007). So würden sich Anforderungen, die in einem angemessenen Passungsverhältnis zu den Kompetenzen der Lernenden stehen, positiv auf die subjektiven Kontrollüberzeugungen und Wertzuschreibungen auswirken, durch die wiederum Gefühle der Langeweile unwahrscheinlicher werden (Pekrun, 2006). Anforderungen, die die Kompetenzen der Lernenden übersteigen oder unterschreiten, würden hingegen zu sehr niedrigen bzw. sehr hohen Kontrollüberzeugungen führen und Langeweile fördern (Goetz & Hall, 2014).

Das Passungsverhältnis zwischen Anforderungen und Kompetenzen lässt sich auf den Sprachgebrauch im Lehr- und Lernkontext übertragen, der sich auf lexikalischer und grammatikalischer Ebene vom alltäglichen Sprachgebrauch unterscheidet und mit erhöhten Anforderungen in der Nutzung einhergeht (Gogolin & Lange, 2011). Für den erfolgreichen Umgang mit diesen sprachlichen Anforderungen benötigen Lernende sprachliche Kompetenzen, die eine rezeptive und produktive Verarbeitung der über die Sprache vermittelten Lerninhalte ermöglichen (Morek & Heller, 2012). Inwieweit sich das Verhältnis dieser beiden Aspekte auf die Langeweile auswirkt, bleibt in der Forschung bislang unberücksichtigt.

Fragestellungen

Basierend auf diesen theoretischen Grundlagen wurde in der vorliegenden Studie untersucht, (1) wie sich die situationale Langeweile über eine ausgewählte Lehr-Lernsituation entwickelt, (2) ob sich heterogene Entwicklungsverläufe situationaler Langeweile identifizieren lassen und (3) ob die Passung sprachlicher Anforderungen und der individuellen sprachlichen Kompetenzen mit der Entwicklung situationaler Langeweile zusammenhängt.

Methode

An der Datenerhebung im Wintersemester 2021/2022 nahmen N=123 Lehramtsstudierende (87.0% weiblich) teil. Die Studierenden wurden zufallsbasiert einer von drei Versuchsbedingungen zugewiesen, in denen ein Erklärvideo zu einem methodischen Sachverhalt in Anlehnung an das

Sprachmodell von Heine et al. (2018) sprachlich variiert wurde. Die einfache Sprachversion richtete sich nach dem Alltagssprachgebrauch, verwendete hierfür vorwiegend eine direkte Ansprache, aktive Verbformen kurze und einfache Sätze. Die schwierige Sprachversion enthielt komplexere Satzstrukturen mit hoher Informationsdichte, Nominalisierungen und Passivkonstruktionen. Die mittelschwere Sprachversion stellt eine Mischform aus den anderen beiden Sprachversionen dar. Anhand eines C-Tests zu Beginn der Datenerhebung wurde die sprachliche Kompetenz der Lernenden erfasst. Die Langeweile wurde zu drei Messzeitpunkten erhoben: vor und nach dem Erklärvideo sowie nach einem Test zum Themenverständnis, der auf das Erklärvideo folgte, um einen Einblick in prozesshafte Veränderungen der Langeweile zu ermöglichen.

Ergebnisse

Bezugnehmend auf die erste Forschungsfrage wurde deutlich, dass die Langeweile über alle Studierenden hinweg im Verlauf der Lehr-Lernsituation von Messzeitpunkt (MZP) 1 zu MZP 2 signifikant abnahm und zu MZP 3 auf diesem Niveau verblieb. Zur zweiten Forschungsfrage zeigten Latent Class Growth Analysen zwei Typen von Studierenden mit unterschiedlicher Entwicklung der Langeweile. Während sie in der einen Gruppe (n=98) im Verlauf signifikant abnahm, stieg die Langeweile in der anderen Gruppe (n=25) über die drei Messzeitpunkte hinweg signifikant an. Im Hinblick auf die dritte Forschungsfrage nach dem Zusammenhang von Passung zwischen sprachlichen Anforderungen des Erklärvideos und den sprachlichen Kompetenzen zeigte sich ein signifikanter Zusammenhang der Klassenzugehörigkeit mit den C-Test-Ergebnissen bei Studierenden, die das sprachlich schwierige Erklärvideo zugeteilt bekommen hatten (n=38). Ein Mittelwertvergleich offenbarte, dass die Studierenden, deren Langeweile anstieg, signifikant schlechtere Ergebnisse im C-Test erzielten als Studierende, deren Langeweile im Verlauf abnahm. Die Ergebnisse geben damit Hinweise darauf, dass nicht-optimale Passungsverhältnisse von Anforderungen und Kompetenzen Langeweile fördern können und werden abschließend im Hinblick auf die Bedeutung der sprachlichen Gestaltung (hoch)schulischer Bildungsprozesse diskutiert.

D-14: Schulischer Wandel in der Migrationsgesellschaft – Schulkultur(en) im Kontext aktueller Fluchtmigration

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* T03 R02 D81

Schulischer Wandel in der Migrationsgesellschaft Migrationsgesellschaft – Schulkultur(en) im Kontext aktueller Fluchtmigration

Chair(s): **Dominique Patrizia Rauch** (PH Ludwigsburg, Germany)
*Diskutant*in(nen):* **Sara Fürstenau** (Universität Hamburg)

Das Symposium stellt erste Ergebnisse aus dem Projekt SchuWaMi vor, das untersucht, wie Schulen in Deutschland auf die vermehrte Aufnahme von geflohenen Kindern und Jugendlichen reagiert haben und welche institutionellen Veränderungsprozesse in diesem Kontext stattgefunden haben und immer noch stattfinden. Zentral ist dabei die Rolle der Schulkulturen und ihre institutionellen Voraussetzungen, die in unterschiedlichen Aspekten einmal quantitativ (Beitrag 2) und einmal qualitativ (Beitrag 3) fokussiert werden. Überzeugungen und Praktiken von Lehrkräften als zentralen professionelle Akteuren in den Klassen und Schulen werden in Beitrag 1 und 2 analysiert, wobei letzterer einen multifokalen Zugang über verschiedene schulische Akteursgruppen wählt. Beitrag 3 befasst sich dagegen vertiefend mit der Positionierung von Schulleitungen.

Der erste Beitrag untersucht Zusammenhänge zwischen formalen und informalen Lerngelegenheiten, Überzeugungen zu Mehrsprachigkeit und dem Lehrkraftverhalten im Unterricht, auf der Basis der Fragebogendaten von 375 Lehrkräften von 30 Schulen. Es zeigt sich, dass Überzeugungen zu Mehrsprachigkeit ein relevanter Bestandteil von Lehrerkompetenzen sind und mit formalen und informalen Lerngelegenheiten einerseits und mit professionellem Verhalten im Unterricht andererseits zusammenhängen. Der Anteil von Lehrkräften mit DaZ Qualifikationen kann dabei zur Erklärung des Zusammenhangs zwischen Überzeugungen und Praktiken auf Schulebene beitragen.

Der zweite Beitrag bezieht neben der Perspektive der Lehrkräfte auch jene der Schulleitungen, Schulsozialarbeiter*innen und Schüler*innen mit ein. Er untersucht Ausprägungen der Haltungen, Praktiken und Erfahrungen der verschiedenen Akteure in Schulen mit kontrastierenden idealen schulkulturellen Entwürfen zum Umgang mit Diversität auf Basis von Fragebogendaten von fünf Schulleitungen, 144 Lehrkräften und 175 Sek I Schüler*innen. In den Analysen zeigen sich charakteristische Schulunterschiede in den individuellen Einstellungen und Praktiken der Schulleitungen und Lehrpersonen, die idealen Entwürfe der Einzelschulen widerspiegeln aber auch Spannungsverhältnisse andeuten. Das Erleben der Schüler*innen mit Migrationshintergrund scheint hiervon wenig abzuhängen: Sie fühlen sich an allen Schulen ähnlich wohl und von ihren Lehrkräften wie Mitschüler*innen anerkannt.

Der dritte Beitrag untersucht, wie sich Schulleitungen zu der vermehrten Aufnahme von Seiteneinsteiger*innen mit Fluchterfahrung positionieren und wie dies die (imaginären) schulkulturellen Entwürfe der Einzelschulen prägt. Datenbasis bilden vier teilstrukturierte narrative Interviews, die mit der dokumentarischen Methode ausgewertet wurden. Es zeigt sich, dass sich Schulentwürfe entlang einer passiven vs. aktiven Rolle mit Bezug auf die Aufnahme und Beschulung von Seiteneinsteiger*innen unterscheiden lassen, wobei aktiv bedeutet, dass Schulleitungen ihre Schule als Institution von hoher Expertise für Migrationsthemen und Mehrsprachigkeit verstehen.

In der Zusammenschau werden die Ergebnisse sowohl mit Bezug auf die Möglichkeiten der gemeinsamen Interpretation der vorliegenden Ergebnisse als auch Limitationen der gewählten methodischen Zugänge diskutiert.

Beiträge des Symposiums

Überzeugungen und Praktiken von Lehrkräften in Bezug auf Mehrsprachigkeit

Dominique Rauch[1], Birgit Becker[2]
[1]PH Ludwigsburg, [2]Universität Frankfurrt

Theoretischer Hintergrund

Etwa ein Drittel aller Schüler*innen spricht eine andere Herkunftssprache als Deutsch, die kaum im regulären Unterricht berücksichtigt wird (Becker-Mrotzek et al., 2012). Schule ist ein Ort, an dem überwiegend monolingual Deutsch interagiert wird (Gogolin, 1994) und ein defizitärer Blick auf die Mehrsprachigkeit von Schüler*innen zu beklagen ist (Borowski, 2021). Gleichzeitig ist es notwendig, dass Lehrkräfte über Kompetenzen im Umgang mit der sprachlichen Heterogenität in Schulklassen verfügen, um mehrsprachigen Schüler*innen höhere gleiche Bildungschancen zu ermöglichen (Fischer & Ehmke, 2019). Ein Aspekt professioneller Kompetenz von Lehrkräften im Kontext heterogener Klassen sind Werthaltungen und Überzeugungen (Hachfeld et al., 2012). Überzeugungen zu Mehrsprachigkeit, als individuelles mentales Konstrukt mit kognitiven und affektiven Anteilen (Skott, 2015), werden nach dem Modell von Lange und Pohlmann-Rother (2020) unter anderem geprägt von der Nutzung formaler und informaler Lernangeboten, wie bspw. DaZ-Weiterbildungen und wirken sich andererseits auf professionelles Verhalten im Unterricht aus.

Forschungsfragen

Die vorliegende Studie untersucht die Zusammenhänge zwischen formalen/ informalen Lerngelegenheiten, Überzeugungen zu Mehrsprachigkeit und dem Lehrkraftverhalten im Unterricht.

Methode

Die Forschungsfrage wird im Rahmen der quantitativen Teilstudie des SchuWaMi-Projektes untersucht. 375 Lehrkräfte von 30 Schulen (N=3 Gymnasien, N=9 Integrierte Gesamtschulen und N=18 Sekundarschulen) aus vier Bundesländern beantworteten den Lehrkräftefragebogen. 70 % der befragten Lehrerinnen waren weiblich, 10 % hatten einen Migrationshintergrund. Die Lehrkräfte hatten im Mittel 13,3 Jahre Berufserfahrung. 77 % haben bereits Seiteneinsteiger*innen im deutschen Bildungssystem unterrichtet. Die Lehrkräfte beantworteten Fragen zur Beschäftigung mit Deutsch als Zweitsprache (DaZ) und Mehrsprachigkeit im Studium, in der Fort- oder Weiterbildung und in der praktischen Arbeit. Überzeugungen zu Mehrsprachigkeit wurden mit sechs Items mit einer sechsstufigen Antwortskala (trifft überhaupt nicht zu – trifft völlig zu) erfasst (Cronbachs Alpha .827; Beispielitem: Mehrsprachigkeit in der Schule hat häufig zur Folge, dass die Förderung der deutschen Sprache bei allen Schüler*innen zu kurz kommt) und der Umgang mit Mehrsprachigkeit im Unterricht mit 10 Items mit vierstufiger Antwortskala erfasst (Beispielitem: Mehrsprachige Schüler*innen dürfen bei Arbeitsaufträgen in Herkunftssprache kommunizieren).

Ergebnisse und ihre Bedeutung

Die Hälfte der teilnehmende Lehrkräfte beschäftigten sich in der praktischen Arbeit mit DaZ (55%) Mehrsprachigkeit (52%), etwa ein Drittel (DaZ: 31%; Mehrsprachigkeit: 26%) haben sich bereits im Studium mit diesen Themen befasst, Fortbildungen wurden seltener genannt (DaZ: 24%; Mehrsprachigkeit: 15%).

Lehrkräfte geben an Praktiken, die auf die Verwendung des Deutschen fokussieren, sehr häufig einzusetzen (bspw. Begriffsklärung), während sie explizit auf Mehrsprachigkeit ausgerichtete Praktiken (bspw. Verwendung mehrsprachige Materialien) eher selten einsetzen. Lehrkräfte, die sich bereits mit DaZ und/oder Mehrsprachigkeit beschäftigt haben, haben positivere Überzeugungen zu Mehrsprachigkeit (rDaZ =.20**; rMehrsprachigkeit =.25**). Auch zwischen einigen Items zu Praktiken im Unterricht und den Überzeugungen zu Mehrsprachigkeit zeigten sich signifikante Zusammenhänge (r zwischen .197** und .415**). Mittlere Korrelationen zeigten sich vor allem für Items, die sich direkt auf das Zulassen oder aktive Einbeziehen von Mehrsprachigkeit beziehen. Auf Schulebene korrelieren Überzeugungen zu Mehrsprachigkeit und Praktiken im Unterricht zu r=.52**. Betrachtet man Einzelschulen, die deutlich von diesem mittleren Zusammenhang abweichen, so lässt sich diese Abweichung über Unterschiede in den jeweiligen Anteilen an Lehrkräften mit DAZ/DAF-Ausbildung erklären.

Die Ergebnisse deuten darauf hin, dass Überzeugungen zu Mehrsprachigkeit ein relevanter Bestandteil von Lehrerkompetenzen sind, die - wie von Lange und Pohlmann-Rother (2020) postuliert - sowohl mit formalen und informalen Lerngelegenheiten als auch mit professionellem Verhalten im Unterricht zusammenhängen. Die Ergebnisse werden vor dem Hintergrund der Limitationen (Querschnitt; Stichprobengröße, Datenstruktur), weiterer Analyseperspektiven (Mediationen) und ihrer Bedeutung für die Praxis diskutiert (Überzeugungen als Dreh- und Angelpunkt pädagogischer Praxis).

Ideale Entwürfe und schulische Praxis des Umgangs mit Diversität in der Schule

Svenja Vieluf
TU Braunschweig

Theoretischer Hintergrund

Der schulische Umgang mit migrationsbedingter Diversität ist geprägt von übergreifenden migrations- und bildungspolitischen Vorgaben, gesamtgesellschaftlichen Verhältnissen, Strukturen und Diskursen. Doch in jeder Einzelschule setzen sich die Akteure je spezifisch handelnd mit diesem Rahmen auseinander. Dabei entstehen einerseits ideale Entwürfe, die u.a. in Schulprogrammen formuliert und auf Schulhomepages präsentiert werden, und andererseits schulspezifische Werte, Normen und Praktiken, die sich in alltäglichen Interaktionen manifestieren. Die symbolische Ordnung der Schulkultur wird als einzelschulspezifisch ausgeformtes Spannungsverhältnis dieser drei Ebenen gefasst. Die jeweilige Passung individueller Haltungen und Routinen zu den schulkulturellen Entwürfen spiegelt sich zudem in innerschulischen Anerkennungsverhältnissen wider (vgl. u.a. Helsper 2008; Helsper, Böhme, Kramer, & Lingkost, 2001).

Forschungsfragen

Die vorliegende Studie zielt darauf ab, Beziehungen zwischen idealen schulkulturellen Entwürfen zum Umgang mit Diversität und individuellen Haltungen und Praktiken der schulischen Akteure zu

beleuchten und außerdem zu untersuchen, ob sich entsprechende Konstellationen auch in den Schulerfahrungen von Schüler*innen mit Migrationshintergrund widerspiegeln.

Methode

Der Forschungsfrage wird im Rahmen der quantitativen Teilstudie des SchuWaMi-Projektes nachgegangen. In den teilnehmenden Schulen haben Schulleitungen, Lehrkräfte Schulsozialarbeiter*innen und Schüler*innen Fragebögen beantwortet, die u.a. Fragen zu ihren Einstellungen zu Diversität, ihrem Umgang mit Mehrsprachigkeit und Leistungsheterogenität in Schule und Unterricht und ihrer Wahrnehmung schulischer Anerkennungsverhältnisse enthielten. Im vorliegenden Beitrag werden die Antworten in fünf Schulen verglichen, die sich in Bezug auf zwei Aspekte ihrer idealen Entwürfe möglichst stark unterscheiden: a) Inwieweit ein konstruktiver Umgang mit kultureller Diversität als Ideal formuliert wird und b) inwieweit ein besonderer Ansatz zum Umgang mit Leistungsheterogenität verfolgt wird. Ausgewählt wurden auf dieser Basis:

- eine IGS, welche langes gemeinsames Lernen realisiert und Wochenplanarbeit sowie die Förderung überfachlicher Kompetenzen fest etabliert hat.

- eine kooperative Gesamtschule, die einen besonderen Schwerpunkt im Bereich der kulturellen Bildung anbietet, Interkulturalität fördert und sich rassismuskritisch positioniert.

- eine Hauptschule, die herkunftssprachlichen Unterricht in verschiedenen Sprachen anbietet und Wert auf ein Empowerment ihrer diversen Schüler*innenschaft legt.

- eine Hauptschule, die Interkulturalität fördert und sich rassismuskritisch positioniert.

- zwei Realschulen, die in Bezug auf den Umgang mit Diversität und Heterogenität keine besonderen Konzepte oder Positionierungen auf der Homepage präsentieren.

In diesen Schulen haben 5 Schulleitungen, 144 Lehrkräfte und 175 Schüler*innen (aus 2 Klassen an der jeweiligen Schule) Fragebögen beantwortet, die u.a. Fragen zu ihren Einstellungen zu Diversität, zu ihrem Umgang mit Mehrsprachigkeit und Leistungsheterogenität in Schule und Unterricht und zu ihrem Wohlbefinden in der Schule und ihrer Wahrnehmung schulischer Anerkennungsverhältnisse enthielten. Für die sechs Schulen wurden jeweils prozentuale Anteile oder Mittelwerte für diese Variablen unter Berücksichtigung der Standardfehler in der Gesamtstichprobe verglichen.

Ergebnisse und Diskussion

In den Analysen zeigen sich charakteristische Schulunterschiede in den individuellen Einstellungen und Praktiken der Schulleitungen und Lehrpersonen, die teilweise die idealen Entwürfe der Einzelschulen widerspiegeln, teilweise aber auch Spannungsverhältnisse andeuten. Insbesondere zeigen sich die Akteure in Schulen, die auf ihrer Homepage die Interkulturalität der Schüler*innenschaft hervorheben und/oder sich rassismuskritisch positionieren, offener für die Nutzung von Herkunftssprachen in der Schule als die Akteure in den anderen Schulen. In der IGS, die Wochenplanarbeit und überfachliches Lernen etabliert hat, berichten die Lehrkräfte zudem besonders häufig Individualisierung und soziales Lernen zu fördern. Doch im Einklang mit der Annahme von Helsper et al. (2008 u.a.), dass sich in Schulen zwar temporär bestimmte Normen und Praktiken als dominant etablieren, diese jedoch umkämpft bleiben, zeigt sich auch in unseren Daten

innerschulische Varianz in diesen Einstellungen und Praktiken. Schließlich spiegeln sich die beobachteten Schulunterschiede wenig im Erleben der Schüler*innen wider: Schüler*innen mit einem Migrationshintergrund fühlen sich an allen Schulen ähnlich wohl und von ihren Lehrkräften wie Mitschüler*innen anerkannt. Die Ergebnisse sollen mit Rückbezug auf die Schulkulturtheorie kritisch diskutiert werden.

Schulkulturelle Entwürfe im Kontext von Fluchtmigration

Anna Laschewski[1], **Patricia Stosic**[2]
[1]Universität Duisburg-Essed, [2]Universität Duisbugr-Essen

Theoretischer Hintergrund

Durch die vermehrte Neuzuwanderung von Kindern und Jugendlichen in den letzten Jahren, bildet die Frage danach, wie Schulen mit der Beschulung sogenannter Seiteneinsteiger*innen umgehen und welche Anschlussmöglichkeiten aber auch Ausschlüsse sowie (transnationale) bildungsbiografische Perspektivierungen und Positionierungen sich für neuzugewanderte Schüler*innen aus den jeweiligen schulpädagogischen und schulorganisatorischen Umgangsweisen eröffnen, ein zentrales erziehungswissenschaftliches Erkenntnisinteresse (vgl. Diehm & Radtke, 1999; aktuell z.B. Dewitz, Terhart & Massumi, 2018; Emmerich, Hormel & Massumi, 2020; Hummrich, 2020). In unserem Forschungsprojekt „SchuWaMi“ sind wir dieser Frage nachgegangen und haben uns theoretisch wie methodologisch auf den Halleschen Schulkulturansatz (vgl. für einen Überblick Böhme, Hummrich & Kramer, 2015) bezogen. Dieser eröffnet die Möglichkeit, in Schulkulturanalysen den einzelschulspezifischen Umgang mit migrationsspezifischer Differenz und Ungleichheit unter Einbezug gesellschaftlicher Bezüge und Rahmenbedingungen zu analysieren. Die Hervorbringung von Schulkultur(en) als ‚symbolische Sinnordnung(en)‘ (Kramer, 2016) wird in diesem Ansatz als eingebettet in die „handelnde Auseinandersetzung der schulischen Akteure mit übergreifenden, bildungspolitischen Vorgaben und Strukturierungen vor dem Hintergrund historischer Rahmenbedingungen und der sozialen Auseinandersetzung [zur] Durchsetzung und Distinktion pluraler kultureller Ordnungen und deren Hierarchisierung“ (Helsper, 2008, S. 66 f.) verstanden. Unterschieden werden hierbei eine imaginäre, eine reale und eine symbolische Ebene: Auf der Ebene der Einzelschule verweist das Imaginäre auf die „idealen pädagogischen Entwürfe“, wobei Schulleitungen eine herausgehobene Funktion zukommt. Das Symbolische markiert „die Ebene der Entfaltung von Interaktionen, Praktiken, Artefakten, Routinen und Arrangements“ (Helsper, 2008, S. 67 f.) und das Reale der Schulkultur hebt ab auf die Rahmenbedingungen schulpädagogischen Handelns.

Fragestellung

In dem vorliegenden Beitrag soll untersucht werden, wie sich Schulen zu der vermehrten Aufnahme von Seiteneinsteiger*innen mit Fluchterfahrung positionieren und wie dies in die (imaginären) schulkulturellen Entwürfe der Einzelschulen mit einspielt. Dabei wird fokussiert, wie die Schulleitungen, als zentrale Repräsentant*innen und Entscheidungsträger*innen ihrer Einzelschule, retrospektiv die Aufnahme- und schulorganisatorischen Prozesse narrativ einbetten.

Methode

Untersucht wird die dargestellte Forschungsfrage an teilstrukturierten narrativen Interviews vierer Schulleiter*innen von Einzelschulen in Hessen, Nord-Rhein-Westfalen und Baden-Württemberg , die im Zuge der qualitativen Teilstudie des BMBF-Projektes „SchuWaMi“ erhoben wurden. Zur Analyse der (imaginären) schulkulturellen Entwürfe und Positionierungen der Schulleitungen für ihre Schulen, wurden die Interviews mit der Dokumentarischen Methode (Nohl, 2017) ausgewertet.

Ergebnisse und Bedeutung

Befragt wurden vier Schulleiter*innen von Haupt-, Gesamt- und Gemeinschaftsschulen. Die ersten Auswertungen zeigen verschiedene Positionierungen der Schulleitungen für Schulen im Hinblick auf die Eingliederung der neuzugewanderten Schüler*innen. Markant ist, dass die Schulentwürfe sich insbesondere entlang einer passiven vs. aktiven Rolle mit Bezug auf die Aufnahme und Beschulung von Seiteneinsteiger*innen unterscheiden lassen. So lassen sich Fälle, bzw. imaginäre Entwürfe, unterscheiden, in denen Schulen sich aktiv mit den Herausforderungen auseinandersetzen, die sich mit der Eingliederung neuzugewanderter Schüler*innen ergeben (haben), sie betrachten sich selbst als Expertinnen für Migrationsthemen und Mehrsprachigkeit und definieren und legitimieren sich gar über den Status als ‚Flüchtlingsschule‘. Auf der anderen Seite finden sich Schulen, die eher passiv bis abwehrend auf die Zuweisung von neuzugewanderten Schüler*innen reagieren und kaum schulkulturelle sowie organisatorische Integrationsleistungen anstreben. Die Ergebnisse zeigen, dass bzw. wie sich einige Schulen die Transformation zu Nutze machen, die im Zuge der vermehrten Aufnahme von neuzugewanderten Schüler*innen stattgefunden hat und immer noch stattfindet. Deutlich wird auch, dass die Legitimierung im lokalen Bildungsraum von Bedeutung für die Selbstpositionierungen und schulkulturellen Entwürfe der Einzelschulen zu sein scheint. Die Ergebnisse sollen detaillierter und mit Rückbezug auf die Schulkulturtheorie kritisch und auch im Hinblick auf bildungspolitisch relevante Erkenntnisse diskutiert werden.

D-15: Bedarfsorientierte Ressourcensteuerung im Mehrebenensystem – durch differenzierte Mittelzuweisung Bildungsungleichheiten verringern?

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* R11 T00 D03

Bedarfsorientierte Ressourcensteuerung im Mehrebenensystem – durch differenzierte Mittelzuweisung Bildungsungleichheiten verringern?

Chair(s): **Norbert Sendzik** (Leibniz-Institut für Bildungsverläufe (LIfBi); Wissenschaftszentrum Berlin für Sozialforschung (WZB)), **Denise Demski** (Ruhr-Universität Bochum (RUB))

*Diskutant*in(nen):* **Horst Weishaupt** (Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF)), **Martina Diedrich** (Institut für Bildungsmonitoring und Qualitätsentwicklung (IfBQ))

Wiederkehrend belegen empirische Befunde, dass neben primären und sekundären Herkunftseffekten (Boudon, 1974) auch Kontextfaktoren auf der Ebene von Lerngruppe, Schule und Sozialraum maßgeblich für Bildungschancen sind. Lernende an Schulen in ungünstiger Lage erzielen regelmäßig geringere Leistungen (Baumert et al., 2006), gehen seltener an ein Gymnasium über und verlassen die Schule häufiger ohne Abschluss (Jeworutzki et al., 2017). Die soziale Segregation insbesondere in städtischen Ballungsräumen scheint dafür von besonderer Relevanz zu sein (Helbig & Jähnen, 2018).

Hier setzt die Grundidee einer bedarfsorientierten Ressourcensteuerung im Schulbereich an, häufig mit Hilfe von sogenannten Sozialindizes: Schulen mit schwierigen sozialen Rahmenbedingungen sollen etwa zusätzliche Lehrerstellen, Schulsozialarbeiter*innen, Budgets und weitere Formen der Mittelzuweisungen erhalten. Die Schüler*innen dieser Schulen sollen hierdurch von einer besseren Förderung, etwa in kleineren Klassen, zusätzlichen Sprachförder- oder Ganztagsangeboten, profitieren. Seit etwa zehn Jahren wird die Idee zunehmend von Bundesländern sowie Kommunen aufgegriffen. Und auch auf Bundesebene hält sie verstärkt Einzug, etwa mit dem im aktuellen Koalitionsvertrag festgeschriebenen *Startchancen*-Programm.

Während international bereits Befunde vorliegen, dass etwa die Art und Weise der differenzierten Mittelzuweisung sowie strukturelle Rahmenbedingungen den angestrebten Abbau von sozialen Bildungsungleichheiten beeinflussen können (z.B. Franck & Nicaise, 2022), liegt hierzu in der BRD kaum Evidenz vor (z.B. Helbig et al., 2022; Sendzik, 2018; Weishaupt & Kemper, 2016). Mit Blick auf die Studien und entlang von Ansätzen der Governanceforschung (z.B. Altrichter et al., 2007; Bartl et al., 2019) ist davon auszugehen, dass Akteur*innen auf unterschiedlichen Systemebenen eigenlogische Adaptionen der Idee einer bedarfsorientierten Mittelzuweisung (auf Basis von Sozialindikatoren) vornehmen, Vorgaben rekontextualisieren und es zu einer Symbolpolitik kommen kann. An dieser Forschungslücke setzt das interdisziplinäre und diskursiv angelegte Symposium an und geht dabei folgenden übergreifenden Fragen nach:

- Wie werden im Mehrebenensystem zusätzliche Mittel an Schulen in sozial schwieriger Lage zugewiesen? Werden die „richtigen“ Schulen adressiert?
- Welche Faktoren beeinflussen eine bedarfsorientierte Ressourcensteuerung auf und zwischen den verschiedenen Ebenen?
- Welche Implikationen für Praxis und Forschung können abgeleitet werden?

Die ersten beiden Beiträge nehmen anhand des Stadtstaates Hamburg sowie des Flächenlandes Nordrhein-Westfalen die Ressourcenverteilung auf Bundeslandebene in den Blick. Dabei wird insbesondere auf Sozialindizes als datenbasierte Verteilungsgrundlage eingegangen. Der dritte Beitrag zeigt auf Basis von systematischen Recherchen und einer Onlinebefragung, inwiefern eine bedarfsorientierte Ressourcenverteilung in deutschen Kommunen erfolgt und wie diese ausgestaltet

ist. Der vierte Beitrag fokussiert die Bundesebene und legt anhand aktueller Programme dar, welche Verteilungsprinzipien sich zum Abbau von Bildungsbarrieren identifizieren lassen. Zwei Diskussionsbeiträge zur wissenschaftlichen und politischen-administrativen Bedeutsamkeit und weiteren Implikationen runden das Symposium ab.

Beiträge des Symposiums

Der Hamburger Sozialindex – Praxisbeispiel für eine bedarfsorientierte Ressourcensteuerung von Schulen

Christian Möller, Klaudia Schulte
Institut für Bildungsmonitoring und Qualitätsentwicklung (IfBQ)

Soziale Ungleichheit ist ein bekanntes Phänomen. In Deutschland besteht in den größeren Städten und Stadtstaaten seit langem ein deutliches Gefälle im Sozialstatus der Bevölkerung (Behörde für Stadtentwicklung und Wohnen, 2020). Dieses Gefälle ist insofern von besonderer bildungspolitischer Bedeutung, als sich besonders in Deutschland ein gleichbleibend hoher Zusammenhang zwischen dem sozialen Hintergrund und den Leistungsergebnissen von Schüler*innen nachweisen lässt (Baumert, Stanat & Watermann, 2006; OECD, 2018; Reiss et al., 2019). Lernende aus sozio-ökonomisch ungünstigen Verhältnissen sind oft doppelt benachteiligt: Sie wachsen häufig in einem Umfeld mangelnder sozialer Unterstützung auf und lernen dann an Schulen, an denen vornehmlich Lernende aus eher deprivierten Umfeldern beschult werden (Schümer, 2004; Groos & Jehles, 2015; Deutsches Jugendinstitut, 2019). Die Schulen, die primär Schüler*innen aus sozio-ökonomisch benachteiligten Verhältnissen beschulen, sind pädagogisch besonders herausgefordert, da sie Lernrückstände aufarbeiten und ein benachteiligtes Umfeld kompensieren müssen (Webs et al., 2018).

Ein bildungspolitisches, national und international eingesetztes Instrument, um diesem Ungleichgewicht oder kumulativen Benachteiligungen (Beierle, Hoch & Reißig, 2019) zu begegnen, ist die sozialindikatorenbasierte Ressourcenallokation für Schulen (Morris-Lange, 2016; Klemm & Kneuper, 2017; Weishaupt, 2016).

In Hamburg gibt es seit 1996 einen Sozialindex für alle staatlichen allgemeinen Schulen. Dieser basiert seit 2006 (vgl. Bos et al. 2006) u.a. auf der Kapitaltheorie von Bourdieu (1982, 1983) und beschreibt die sozio-ökonomische Zusammensetzung der Schülerschaft an allen allgemeinen Hamburger Schulen auf einer Skala von 1 bis 6. (vgl. Schulte, Hartig & Pietsch, 2014).

Schulen mit niedrigen Indizes können beispielsweise kleinere Klassen einrichten oder erhalten mehr Mittel für Inklusion sowie Sprachförderung. Dadurch wird sichergestellt, dass Schüler*innen mit geringeren Bildungschancen in der Schule zusätzlich gefördert werden können.

Darüber hinaus wird der Sozialindex für die Berechnung von Vergleichswerten für schulbezogene Ergebnisrückmeldungen im Rahmen von KERMIT oder im Kontext der Bildungsberichterstattung verwendet (vgl. Schulte, Hartig & Pietsch, 2014).

Das methodische Vorgehen bei der Berechnung hat sich über die Jahre stark verändert. Während für die Aktualisierung 2013 eine Schüler*innen- und Elternbefragung mit über 35.000 verteilten

Fragebögen als zentrale Grundlage diente (vgl. Schulte, Hartig & Pietsch, 2014), wurde 2021 ausschließlich auf amtlich verfügbare Daten zurückgegriffen.

Nach einem sorgfältigen Auswahlprozess wurden acht Kriterien spezifiziert, die entweder auf Ebene der Schule (Familiensprache, Sonderpädagogischer Förderbedarf in den Bereichen Lernen, Sprache oder soziale und emotionale Entwicklung, Leistungen des Bildungs- und Teilhabepakets) oder auf Ebene der statistischen Gebiete vorliegen, in denen die Schüler*innen einer Schule leben (Absolvent*innen mit allgemeiner Hochschulreife, nicht erwerbsfähige Hilfebedürftige unter 15 Jahren, Hilfen zur Erziehung, Arbeitslosigkeit, Wahlbeteiligung). In die Berechnung gingen die Daten aller Schüler*innen der Jahrgänge 1 bis 10 ein, sofern sie Regelklassen besuchen. Die Berechnung des Sozialindex erfolgte mithilfe einer konfirmatorischen Faktorenanalyse.

Das spezifizierte Ein-Faktor-Modell weist dabei einen sehr guten Gesamtfit ($\chi^2 = 5.243$, $df = 3$, $p = .155$; RMSEA = .049, CFI = 0.999; TLI = 0.990) und hohe Zusammenhänge zum vorherigen Index auf (vgl. Schulte et al., i. Vorb.). Darüber hinaus zeigen sich die bekannten hohen Zusammenhänge zwischen dem sozio-ökonomischen Hintergrund und Schüler*innenleistungen.

Ungeachtet der für die Konstruktion eines Sozialindex hilfreichen Korrelationsanalysen, die einen sehr hohen Zusammenhang zwischen Sozialindex und Schüler*innenleistungen zeigen, wird an dieser Stelle die bildungspolitische Dimension deutlich: Trotz hohen Ressourceneinsatzes deutet in den Analysen der Schüler*innenleistungen auf Systemebene wenig darauf hin, dass sich der Zusammenhang zwischen dem sozio-ökonomischen Hintergrund und der Schüler*innenleistungen in Hamburg abschwächt. Das schließt nicht aus, dass es Effekte in anderen Bereichen gibt, die jenseits von Schüler*innenleistungen liegen. Eine differenzierte Evaluation der Maßnahmen, die über Schüler*innenleistungen hinausgeht und auch emotionale und soziale Effekte bei Schüler*innen und deren Lehrkräften thematisiert, wäre aus unserer Sicht daher wünschenswert. Überlegungen, wie eine solche Evaluation aussehen könnte sowie erste Analyseansätze werden im Rahmen des Beitrags vorgestellt.

Der Sozialindex für Schulen in NRW. Steuerungsrelevantes Wissen über einzelne Schulen

Jörg-Peter Schräpler, Sebastian Jeworutzki
Ruhr-Universität Bochum (RUB)

Mit einem Sozialindex sollen die spezifischen Gegebenheiten vor Ort an der Schule berücksichtigt werden. Als Steuerungsinstrument soll er eine gezielte Förderung dieser Schulen ermöglichen und den sozialen, individuellen und kollektiven Benachteiligungen der Schülerinnen und Schüler einer Schule Rechnung tragen. Einen solchen Sozialindex gibt es mittlerweile in Nordrhein-Westfalen.

2020 wurde in Nordrhein-Westfalen (NRW) ein schulscharfer Sozialindex von den Autoren im Auftrag des Ministeriums für Schule und Bildung des Landes NRW (MSB) und in Zusammenarbeit mit dem pädagogischen Landesinstitut (QUA-LiS NRW) entwickelt. Dieser wurde 2021 erstmalig von der Landesregierung zur Ressourcenzuweisung an Schulen in NRW eingesetzt (vgl. Schräpler & Jeworutzki, 2021).

Ziel des Index ist es, den Unterstützungsbedarf von Schulen zu identifizieren, der sich aufgrund der sozialen Zusammensetzung der Schülerinnen und Schüler einer Schule ergibt. Das Verfahren in

NRW zur Konstruktion eines Sozialindex nutzt amtliche Schuldaten und einheitliche kleinräumige Daten, die flächendeckend sowie kostengünstig verfügbar sind. Die soziale Zusammensetzung der Schülerinnen und Schüler der Schulen wird über folgende Indikatoren abgebildet und zu einem Indexwert zusammengefasst.

- Kinder- und Jugendarmut: Derzeit gibt es in der amtlichen Schulstatistik NRWs und bei anderen möglichen Datengebern keine landesweit einheitlichen Daten zum Anteil von Schülerinnen und Schülern in Armut an einzelnen Schulen. Es hat sich jedoch gezeigt, dass durch räumliche statistische Verfahren eine gute Approximation möglich ist (vgl. z.B. Jeworutzki et al., 2017, Kraus, Weishaupt & Hosenfeld, 2021). Der verwendete Sozialraumindikator basiert auf der Dichte der SGB II-Quote der Minderjährigen im geschätzten Einzugsgebiet der Grundschulen (näheres hierzu in Schräpler & Jeworutzki, 2016 & 2021). Bei den Schulen der Sekundarstufe wird dieser Indikator aus dem Sozialraumindikatoren der Herkunftsgrundschulen der Schülerinnen und Schülern abgeleitet.

- Schülerinnen und Schüler mit überwiegend nichtdeutscher Familiensprache: Der schulische Erfolg von Schülerinnen und Schülern setzt umfangreiche sprachliche Kompetenzen voraus. Viele Aufgabenstellungen stehen in einem sprachlichen Kontext bzw. sind mit einem der Sprache eng zusammenhängenden kulturellen Kontext verbunden. Insofern steht auch das Niveau der deutschen Sprachkenntnisse in einem engen Zusammenhang mit den Schulleistungen von Kindern und Jugendlichen mit Migrationshintergrund und wird indirekt auch von der im Elternhaus gesprochenen Sprache beeinflusst. Der Anteil an Schülerinnen und Schülern mit nichtdeutscher Familiensprache ist daher ein wichtiger Indikator für die soziale Zusammensetzung der Schülerschaft an Grund- als auch Sekundarschulen.

- Schülerinnen und Schüler mit eigenem Zuzug aus dem Ausland: Die eigene Migrationserfahrung von Schülerinnen und Schülern ist zwar eng mit dem Spracherwerb verbunden, bedingt jedoch unter Umständen weitere Ursachen individueller Benachteiligung. So ist mit ihr oft der Verlust oder eine Entwertung von vorhandenem Bildungskapital verbunden.

- Schülerinnen und Schüler mit sonderpädagogischem Förderschwerpunkt LSE: Das gemeinsame Lernen von Schülerinnen und Schülern mit und ohne sonderpädagogische Förderung stellt für die betroffenen Schulen eine zusätzliche Herausforderung dar. Gerade Schülerinnen und Schüler mit den Förderbedarfen Lernen, Sprache und emotionale und soziale Entwicklung (LSE) kommen häufig aus ökonomisch prekären und sozial sowie auch gesundheitlich belasteten Familien. Insofern wird die Inklusion von Förderschülerinnen und Förderschülern zu einem großen Teil von Schulen in schwieriger Lage geleistet. Die doppelte Herausforderung bei Schulen durch eine Lage in sozial benachteiligten Quartieren und umfangreicheren Inklusionsaufgaben wird bei der Sozialindexkonstruktion daher mitberücksichtigt.

Auf Basis der vier Indikatoren werden für die Grundschulen als auch für die weiterführenden Schulen jeweils eine konfirmatorische Faktorenanalyse durchgeführt und ein Faktormodell geschätzt. Bei der Evaluierung des Index auf Basis der Lernstandsergebnisse VERA 3 und 8 zeigt sich eine hohe Varianzaufklärung. Der Sozialindex wurde für alle öffentlichen allgemeinen allgemeinbildenden Schulen (Primar- und Sekundarbereich) berechnet.

Die Fragen, die mit dem praktischen Einsatz eines Sozialindex verbunden sind, sind weniger Gegenstand wissenschaftlicher Betrachtungen, sondern politischer Natur. Über Art und Umfang einer sozialindexgesteuerten Ressourcenzuweisung entscheiden vor allem die politisch Verantwortlichen.

Ungleiches ungleich behandeln. Empirische Befunde zur bedarfsorientierten Ressourcensteuerung in deutschen Kommunen

Norbert Sendzik[1], Denise Demski[2], Gabriele Bellenberg[2], Marcel Helbig[1]
[1]Leibniz-Institut für Bildungsverläufe (LIfBi); Wissenschaftszentrum Berlin für Sozialforschung (WZB), [2]Ruhr-Universität Bochum (RUB)

Herkunftsbedingte Disparitäten im Bildungssystem sind für unterschiedliche Bildungsetappen und im Zeitverlauf international und national vielfach empirisch belegt, wobei für ihre Erklärung neben primären und sekundären Herkunftseffekten (Boudon, 1974) auch dem Sozialraum und sozialen Segregationstendenzen Relevanz zukommt (z.B. Baumert et al., 2006; Helbig, 2010; Makles et al., 2019; Jeworutzki et al., 2017; Stubbe et al., 2017). Eine bedarfsorientierte Ressourcensteuerung, im Zuge derer Bildungseinrichtungen an herausfordernden Standorten und mit einem hohen Anteil von Kindern und Jugendlichen aus benachteiligten Familien zusätzliche Ressourcen erhalten, wird als (potenzielles) Mittel zum Abbau von Bildungsbarrieren angesehen und seit einigen Jahren auch auf kommunaler Ebene verstärkt eingesetzt (Beierle, Hoch, & Reißig, 2019; Sendzik, 2018; Möller & Bellenberg, 2018; Groos & Jehles, 2015). Im Mehrebenensystem kommt Kommunen dabei eine Doppelrolle hinsichtlich der Mittelzuweisung zu, da sie sowohl über eigene Ressourcen verfügen, daneben z.T. aber auch für die konkrete Verteilung von Bundes- oder Landesmitteln verantwortlich sind. Ansätzen der Educational Governance (z. B. Altrichter et al., 2007) folgend ist dabei von Rekontextualisierungen (Fend, 2008) und eigenlogischen Adaptionen auszugehen und Verteilungsentscheidungen können als Ergebnis sozialer Konstruktionsprozesse (Berger & Luckmann, 2004) aufgefasst werden. In der Folge ist eine beträchtliche Variation hinsichtlich Umfang und Art der zugewiesenen Ressourcen (Stellen, Finanzmittel etc.) sowie der Verteilungsgrundsätze (z. B. Zuweisung aufgrund von Sozialindizes oder sonstigen Daten, Anträgen oder Expertise von Personen in Schulträgerschaft/Bildungsverwaltung) zwischen Kommunen zu erwarten.

Internationale Befunde deuten darauf hin, dass der Ressourcenumfang sowie die Art und Weise, wie diese zugewiesen und genutzt werden (sollen), die Wirkungen einer bedarfsorientierten Ressourcensteuerung stark beeinflussen (z.B. Franck & Nicaise, 2022). Für Deutschland liegt noch kaum Evidenz in Bezug auf Ausmaß, Ausgestaltungen und Wirkungen einer bedarfsorientierten Ressourcensteuerung vor und nach unserer Kenntnis existiert keine systematische Übersicht über die konkrete Umsetzung einer derartigen Mittelzuweisung auf kommunaler Ebene. Leitende Forschungsfragen des Beitrags sind daher:

- Welche Kommunen in Deutschland haben bereits eine bedarfsorientierte Ressourcensteuerung im Schulbereich implementiert?

- Wie wird eine bedarfsorientierte Ressourcensteuerung in Kommunen umgesetzt (z.B. Umfang, Verteilungsprinzipien)?

Zur Beantwortung dieser Fragen wird im BMBF-Verbundprojekt „ABBAUBAR“ (Förderkennzeichen 01JB2103A bzw. 01JB2103B; Laufzeit 04/2022–03/2025) auf unterschiedliche Datenquellen

zurückgegriffen. Im Vortrag werden zum einen Ergebnisse einer aktuell durchgeführten systematischen Dokumentenanalyse präsentiert. Auf Basis der Sichtung und kriteriengeleiteten Codierung von Bildungsberichten, Internetauftritten und Dokumenten in Ratsinformationssystemen sowie der Anfertigung von Fallportraits wird dabei für deutsche Kommunen, die bisher mindestens einen Bildungsbericht veröffentlicht haben, eruiert, ob eine bedarfsorientierte Ressourcensteuerung implementiert wurde und wie diese ausgestaltet ist. Zum anderen wird auf Befunde einer derzeit laufenden Online-Befragung von Verantwortlichen für die kommunale Schulträgerschaft zurückgegriffen, in der die Herkunft von Personal-, Finanz- und Sachmitteln ebenso erfragt werden wie Modalitäten der Zuweisung/Verteilung und Einstellungen zu einer bedarfsorientierten Ressourcensteuerung.

Erste Ergebnisse deuten bereits zum jetzigen Zeitpunkt darauf hin, dass sich die in den Dokumenten beschrieben bzw. in der Befragung berichteten Umsetzungen einer bedarfsorientierten Ressourcensteuerung zwischen Kommunen deutlich unterscheiden. Unterschiede lassen sich dabei sowohl hinsichtlich der Art der Mittel (Personalstellen bzw. Budgets zur freien oder zweckgebundenen Verwendung) identifizieren als auch in Bezug auf die Verteilungsgrundlage (antragsbasiert bzw. datenbasiert, z.B. durch die Verwendung von Sozialindizes). Neben vorliegenden landesweiten Sozialindizes nutzen Kommunen dabei z.T. auch selbstentwickelte Indizes, die erstere ersetzen oder ergänzen. Im Vortrag werden deskriptive Statistiken präsentiert, Einblicke in die unterschiedlichen kommunalen Umsetzungen gegeben und ein erster Vorschlag für eine Typologie dargestellt; zudem werden die Befunde vor dem Hintergrund internationaler Erfahrungen und aktueller Entwicklungen im deutschen Bildungssystem kritisch eingeordnet und Forschungslücken identifiziert. Die zu erwartenden Projektbefunde können wichtige Ergebnisse in einem bisher vernachlässigten Forschungsfeld liefern und zur Weiterentwicklung einer bedarfsorientierten Ressourcensteuerung in den Kommunen beitragen.

Ein Multipler Benachteiligungsindex (MBI) als Alternative zum Königsteiner Schlüssel für die Verteilung von Bundesmitteln im Schulbereich

Detlef Fickermann
Wuppertaler Institut für bildungsökonomische Forschung (WIP); Wissenschaftszentrum Berlin für Sozialforschung (WZB)

Der Königsteiner Schlüssel wurde ursprünglich zur Finanzierung von außeruniversitären Forschungseinrichtungen eingeführt. Er richtet sich zu zwei Dritteln nach dem Steueraufkommen des Landes und zu einem Drittel nach der Bevölkerungszahl. Es sollten sich die Länder, die ein höheres Steueraufkommen haben, mit einem höheren Anteil an der Finanzierung von Forschungseinrichtungen beteiligen (GWK, 2021, WD, 2020). Er wird in der Zwischenzeit aber nicht mehr nur für seinen ursprünglichen Zweck eingesetzt, sondern auch zum Beispiel für die Verteilung von nach Deutschland Geflüchteten auf die Bundesländer und für eine ganze Reihe weiterer Anwendungen (für weitere Anwendungen siehe WD, 2020, S. 9).

Er wird auch für zweckgebundene Zuwendungen des Bundes an die Länder, beispielsweise im Schulbereich auf der Basis des Artikels 104c Grundgesetz, verwendet. Die mit ihm ursprünglich verbundene Intention, die Länder mit einem höheren Steueraufkommen tragen größere Finanzierungslasten, wird geradezu ins Gegenteil verkehrt, wenn nun diese Länder den größten Teil

der Zuwendungen erhalten. Das ist nichts anderes als das „Matthäus-Prinzip“ in Reinform: Wer hat, dem wird gegeben.

Bei einer Anwendung des Königsteiner Schlüssels für die Verteilung von Bundesmitteln an die Länder wird zum Beispiel vernachlässigt, dass die Altersstruktur der Bevölkerungen in den Ländern ganz unterschiedlich ist. Es gibt Länder mit einem vergleichsweise geringen Anteil an Kindern und Jugendlichen im schulpflichtigen Alter. Diese Länder bekommen dann überproportional viel Geld, wenn man es auf eine einzelne Schülerin bzw. einen einzelnen Schüler bezieht. Zudem bleiben sozio-ökonomische Aspekte unberücksichtigt.

Das Ermessen des Bundes bei der Gewährung von Finanz- bzw. Investitionshilfen ist Einschränkungen unterworfen. Es bedarf einer bundeseinheitlichen, willkürfreien, transparenten und nachvollziehbaren und dem Gebot interkommunaler Gleichbehandlung entsprechenden Regelung. Eine Verteilungsentscheidung bedarf zudem einer sachlichen und rational nachvollziehbaren Begründung mit Bezug auf die durch die Förderung verfolgte Zwecksetzung. Alternativen zum Königsteiner Schlüssel sind ohne weiteres denkbar, wenn diese Anforderungen erfüllt sind und darüber ein hinreichender politischer Konsens zwischen den Beteiligten besteht.

Mit Bezug auf Erfahrungen aus England (vgl. DCLG, 2015a, 2015b) und Österreich (vgl. Heintel et al., 2017) haben Fickermann, Schräpler, Weishaupt und Füssel (2022) einen Multiplen Benachteiligungsindex (MBI) für die Verteilung von Bundesmitteln an die Länder im Schulbereich entwickelt, der aus neun Indikatoren besteht, die zu vier Dimensionen zusammengefasst sind. Der entwickelte MBI kann zur Berechnung von Verteilungsschlüsseln genutzt werden, die die Anforderungen, denen Verteilungsschlüssel gerecht werden müssen, vollumfänglich erfüllen.

MBI-basierte Verteilungsschlüssel können auf der Einwohnerzahl eines Landes basieren oder z. B. auch auf der Zahl der Schüler*innen in bestimmten Bildungsbereichen oder aber auch z. B. für Kinder im Kita-Alter.

Im ersten Teil des Vortrages werden der MBI und daraus abgeleitete Verteilungsschlüssel vorgestellt und mit Mittelverteilungen nach dem Königsteiner Schlüssel oder – wie beim Aktionsprogramm „Aufholen nach Corona“ – nach Umsatzsteueranteilen verglichen.

Bund und Länder verhandeln derzeit über das von den Koalitionsparteien in ihrem Koalitionsvertrag vereinbarte Programm Startchancen. Wie das Programm konkret inhaltlich ausgestaltet werden soll, ist noch nicht geklärt. Ebenso ist auch noch offen, wie hoch das Finanzvolumen sein wird und wie das Programm im Rahmen der Möglichkeiten des Bundes unter Beachtung der föderalen Zuständigkeit der Länder für die Schulen finanziert werden soll. Offen ist auch noch, wie die für das Programm vorgesehenen 4.000 Schulen für das Chancenbudget und die 4.000 Schulen für zusätzliche Schulsozialarbeiter*innenstellen ausgewählt werden sollen. Im zweiten Teil des Vortrages wird ein von Fickermann, Helbig, Schräpler und Weishaupt (in Vorbr.) vorgeschlagenes MBI-basiertes Verfahren zur länderübergreifenden Auswahl der Schulen und die zu schaffenden Voraussetzungen für dessen Anwendung vor- und zur Diskussion gestellt.

D-16: Digitale Kompetenzen und Lehrerbildung

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* T03 R03 D89

Digitale Kompetenzen und Lehrerbildung

Chair(s): **Raphaela Porsch** (Otto-von-Guericke Universität Magdeburg, Deutschland), **Charlott Rubach** (Universität Rostock, Deutschland)
*Diskutant*in(nen):* **Andreas Lachner** (Eberhard Karls Universität Tübingen)

Obwohl der Einsatz digitaler Medien in schulischen Lehr-Lern-Prozessen bereits seit mehreren Jahren auf nationaler Ebene (KMK, 2017, 2021) und an vielen Hochschulen in Curricula verankert wurde, ist die Implementation zur Vermittlung digitaler Kompetenzen in der Lehrerbildung bislang noch nicht strukturiert und flächendeckend erfolgt (Porsch, 2021). Neben fehlenden Konzepten zur effektiven Vermittlung ist trotz der Prominenz des Themas wenig über die Ausprägung digitaler Kompetenzen bei angehenden Lehrkräften bekannt. Vor diesem Hintergrund stellt das Symposium drei aktuelle Forschungsarbeiten vor, deren Befunde anschließend diskutiert werden sollen. Im ersten Beitrag wird der Frage nachgegangen, inwieweit Unterschiede in berufsbezogenen digitalen Kompetenzüberzeugungen zwischen Lehramtsstudierenden, Referendar*innen und Lehrkräften vorliegen. Grundlage der Studie mit 1386 Teilnehmer*innen ist ein Instrument, welches auf Grundlage des DigCompEdu-Rahmenmodells (Redecker & Punie, 2017) entwickelt wurde, die Unterscheidung in verschiedene Dimensionen bzw. Teilkompetenzen erlaubt und somit differenzierte Ableitungen der Befunde für die Lehrerbildung in allen drei Phasen ermöglicht werden. Der zweite Beitrag stellt Ergebnisse zu einem Projekt vor, welches die Förderung digitaler Kompetenzen bei Lehrkräften und Lehramtsstudierenden durch sog. Reverse Mentoring als Ziel hat. Dazu entwickelten Lehramtsstudierende, von denen angenommen wird, dass sie über einen sicheren Umgang mit digitalen Medien verfügen, gemeinsam mit Lehrkräften, die eine hohe Expertise im Unterrichten besitzen, Ideen für einen effektiven Einsatz digitaler Medien im Unterricht. Neben der Vorstellung des didaktischen Konzepts, werden Ergebnisse der Begleitforschung vorgestellt. Im dritten Beitrag werden ebenfalls Ergebnisse von zwei Interventionen vorgestellt. Mithilfe sog. Utility-Value Interventionen sollte die TPACK Wissensintegration im Lehramtsstudium unterstützt werden. Hintergrund ist die Annahme, dass Lehrkräfte über technologischen, pädagogischen und fachlichen Inhalten verfügen sollten, d.h. technological-pedagogical-content-knowledge besitzen wie es das TPACK-Modell postuliert (Mishra & Koehler, 2006). Ergänzend wurden die Konzeption der zwei präsentierten Studien in einem experimentellen Design durch Annahmen der Motivationspsychologie geleitet. Bekannt ist, dass Nützlichkeitsüberzeugungen bei Lernenden durch Utility-Value Interventionen beeinflussbar sind. Die Ergebnisse der Studien verweisen vor allem auf Unterschiede in den Lernprozessen und weniger in der Lernwirksamkeit, so dass u.a. Vorschläge für Anschlussstudien zur Diskussion gestellt werden. Das Symposium möchte einen Beitrag zur Beantwortung der folgenden zentralen Fragen geben: Wie kann man digitale Kompetenzen bei angehenden Lehrkräften objektiv und reliabel messen und in welchem Umfang besitzen sie Kompetenzen für den Einsatz in Lehr-Lernsituationen? Wie lässt sich die Förderung digitaler Kompetenzen in der Lehrerbildung didaktisch realisieren und wie kann man diese Konzepte angemessen evaluieren? Das Symposium schließt mit einem Diskussionsbeitrag durch einen Erziehungswissenschaftler, der durch seinen Schwerpunkt im Bereich Lehren und Lernen mit digitalen Medien ausgewiesen ist.

Unterschiede in berufsbezogenen digitalen Kompetenzüberzeugungen zwischen Lehramtsstudierenden, Referendar*innen und Lehrkräften

Jennifer Quast[1], Charlott Rubach[2], Raphaela Porsch[1]
[1]Otto-von-Guericke Universität Magdeburg, Deutschland, [2]Universität Rostock

Die fortschreitende Digitalisierung der Gesellschaft und nicht zuletzt die flächendeckende Einführung von Distanzunterricht im Zuge der Corona-Pandemie unterstreichen die Bedeutung der Kompetenzen von Lehrkräften im Umgang mit digitalen Medien (Sepúlveda, 2020). Nach der Erwartungs-Wert-Theorie (Eccles et al., 1983) beeinflussen Überzeugungen über die eigenen Fähigkeiten zur Bewältigung einer spezifischen Aufgabe das Verhalten einer Person. Für die effektive Integration digitaler Medien in den Unterricht benötigen Lehrkräfte demnach digitale Kompetenzen und entsprechende Kompetenzüberzeugungen (Guggemos & Seufert, 2021). Digitale Kompetenzüberzeugungen lassen sich in berufsunabhängige und berufsbezogene Überzeugungen unterscheiden (Krumsvik, 2014). Mehrere Studien verweisen darauf, dass berufsbezogene digitale Kompetenzüberzeugungen von Lehrkräften mit deren Nutzung digitaler Medien im Unterricht zusammenhängen (Guggemos & Seufert, 2021; Hatlevik & Hatlevik, 2018; Kaarakainen & Saikkonen, 2021). Die Entwicklung und Förderung berufsbezogener digitaler Kompetenzüberzeugungen stellt daher eine wichtige Aufgabe der Lehrkräftebildung dar, die sich über alle drei Phasen – vom Studium über den Vorbereitungsdienst bis hin zur Weiterbildung im Beruf – erstreckt (Eickelmann & Drossel, 2020). Nach unserer Kenntnis hat jedoch bislang keine Studie die berufsbezogenen digitalen Kompetenzüberzeugungen von Lehramtsstudierenden, Lehramtsreferendar*innen und Lehrkräften (nachfolgend als „(angehende) Lehrkräfte“ zusammengefasst) gemeinsam untersucht und verglichen. Daher widmet sich dieser Beitrag der folgenden Fragestellung: Inwiefern unterscheiden sich Lehramtsstudierende, Lehramtsreferendar*innen und Lehrkräfte hinsichtlich ihrer berufsbezogenen digitalen Kompetenzüberzeugungen?

Für die Untersuchung wurden die Daten von N = 1.386 (angehenden) Lehrkräften aus Deutschland verwendet, die zwischen Mai und Juni 2020 an einer 20-minütigen Online-Befragung teilgenommen haben. Die Stichprobe setzt sich aus n = 524 Lehramtsstudierenden (37,8% der Stichprobe, davon 74,8% weiblich), n = 164 Lehramtsreferendar*innen (11,8% der Stichprobe, davon 72,8% weiblich) und n = 698 Lehrkräften (50,4% der Stichprobe, davon 72,0% weiblich) zusammen.

Die berufsbezogenen digitalen Kompetenzüberzeugungen wurden mithilfe eines selbstentwickelten Instruments auf Grundlage des DigCompEdu-Rahmenmodells (Redecker & Punie, 2017) erfasst (Quast et al., eingereicht). Das Instrument umfasst 33 Items in sieben Skalen: Administration und Weiterbildung, Unterrichtsplanung, Lehren und Unterstützung von Schüler*innen, Evaluation, Schüler*innenorientierung, Förderung der Digitalen Kompetenz der Schüler*innen sowie Datenschutz und Urheberrecht. Die Antworten wurden auf einer 5-stufigen Likertskala (1 = gar nicht kompetent bis 5 = sehr kompetent) erfasst.

Zur Untersuchung der Fragestellung wurde ein Mehrgruppenvergleich in Mplus 8 durchgeführt. Die Analyse zeigt, dass zwischen den Gruppen skalare Messinvarianz gegeben ist. Signifikante Mittelwertsunterschiede lassen sich in unterschiedlichem Ausmaß zwischen allen drei Gruppen feststellen: Lehramtsstudierende berichten im Vergleich zu Lehramtsrefendar*innen in vier von sieben

Kompetenzbereichen niedrigere berufsbezogene digitale Kompetenzüberzeugungen. Verglichen mit Lehrkräften geben Lehramtsstudierende in drei Kompetenzbereichen niedrigere, in einem Kompetenzbereich jedoch höhere berufsbezogene digitale Kompetenzüberzeugungen an. Der Vergleich von Lehramtsreferendar*innen und Lehrkräften zeigt, dass Lehramtsreferendar*innen in drei Kompetenzbereichen höhere berufsbezogene digitale Kompetenzüberzeugungen berichten. Insgesamt geben Lehramtsreferendar*innen im Vergleich der drei Gruppen in der Tendenz die höchsten und Lehramtsstudierende die niedrigsten berufsbezogenen digitalen Kompetenzüberzeugungen an.

Überraschend sind die Ergebnisse der Vergleiche zwischen Lehrkräften und einerseits Lehramtsreferendar*innen sowie andererseits Lehramtsstudierenden. Zwar verfügen Lehrkräfte gegenüber den beiden anderen Gruppen, die sich formal noch in Ausbildung befinden, über die längste Berufserfahrung, jedoch unterscheiden sich Lehrkräfte in vier Kompetenzbereichen nicht signifikant von Lehramtsreferendar*innen und in drei Kompetenzbereichen nicht signifikant von Lehramtsstudierenden. Darüber hinaus berichten Lehramtsreferendar*innen in vier Kompetenzbereichen und Lehramtsstudierende in einem Kompetenzbereich sogar höhere Kompetenzüberzeugungen als Lehrkräfte. Die Ergebnisse verweisen somit auf Handlungsbedarf in der Lehrkräftebildung hinsichtlich des kompetenten Umgangs mit digitalen Medien in Schule und Unterricht. Die Entwicklung berufsbezogener digitaler Kompetenzen sollte bereits während des Studiums stärker gefördert werden. Darüber hinaus stellt die Weiterentwicklung berufsbezogener digitaler Kompetenzen von Lehrkräften eine wichtige Aufgabe der dritten Phase der Lehrerbildung dar.

Förderung digitaler Kompetenzen bei Lehrkräften und Lehramtsstudierenden durch Reverse Mentoring

Victoria Bertram, Charlotte Dignath, Franziska Baier-Mosch, Kunter Mareike
DIPF

Digitale Medien bieten viele Möglichkeiten, um Unterricht innovativ und lernförderlich zu gestalten. Lehrkräfte weisen jedoch häufig Lücken in ihrem Wissen über digitale Medien auf und sind der Nutzung im Unterricht eher negativ eingestellt, sodass diese nur selten effektiv zum Einsatz kommen (Bos et al., 2014). Im Vergleich dazu zeigen Lehramtsstudierende häufig positivere Einstellungen zu digitalen Medien und sind vertrauter in ihrer Anwendung, verfügen jedoch nicht über die notwendige Unterrichtserfahrung, um diese effektiv im Unterricht einzusetzen (Beyerbach et al., 2001; Jimoyannis & Komis, 2007). Die Methode Reverse Mentoring ermöglicht den Austausch der unterschiedlichen Expertise von Lehramtsstudierenden und Lehrkräften, indem traditionelle Mentoringstrukturen umgedreht werden: Lehramtsstudierende, die einen sicheren Umgang mit digitalen Medien mitbringen, geben ihr Wissen an erfahrene Lehrkräfte weiter. Lehrkräfte teilen wiederum ihre unterrichtspraktischen Erfahrungen mit den Lehramtsstudierenden. Gemeinsam entwickeln beide Gruppen Ideen für einen effektiven Einsatz digitaler Medien im Unterricht.

Reverse Mentoring hat sich als innovative Lehrmethode bereits in Organisationen (Graf & Edelkraut, 2017), Schulen (Zauchner-Studnicka, 2017) und in der Lehrkräftebildung (Singer & Maher, 2007) als vielversprechend erwiesen, um digitale Kompetenzen zu fördern. Bislang wurde sie jedoch noch nicht

systematisch in quantitativen Studien evaluiert. Unser Beitrag zielt darauf ab, diese Forschungslücke zu schließen und die Effektivität von Reverse Mentoring in einer quasi-experimentellen Interventionsstudie zu untersuchen und mit traditionellem kooperativem Lernen zwischen Lehramtsstudierenden in Hochschulseminaren und Lehrkräften in Fortbildungen zur beruflichen Weiterbildung zu vergleichen. Als abhängige Variablen werden das selbstberichtete technologisch-pädagogische Wissen sowie positive Überzeugungen zum Einsatz digitaler Medien im Unterricht untersucht.

Lehramtsstudierende nahmen an einem Blockseminar mit fünf Sitzungen á 90 Minuten teil, in denen didaktische Funktionen und Potentiale digitaler Medien und der Umgang mit diesen vermittelt wurde. Darauf folgten zwei Workshoptage, an denen Lehramtsstudierende in Tandems untereinander (KG1) oder in Tandems mit einer Lehrkraft (Reverse Mentoring; EG) gemeinsam Ideen zum Einsatz digitaler Medien entwickelten. Lehrkräfte der Kontrollbedingung (KG2) nahmen an einer zweitägigen Fortbildungsveranstaltung teil, in der kooperativ gearbeitet wurde und eine Dozentin Wissen über digitale Medien vermittelte. 36 Lehramtsstudierende der traditionellen Hochschulseminare (KG1) und 19 Lehrkräfte der Fortbildungsveranstaltungen (KG2) nahmen an begleitenden Erhebungen vor und nach den Workshoptagen teil. In einer ersten Pilotierung wurde die Reverse Mentoring-Bedingung mit 14 Tandems mit je einer Lehrkraft und einer/m Studierenden durchgeführt. Das selbstberichtete technologisch-pädagogische Wissen wurde mit neun Items auf einer sechsstufigen Likert-Skala in Anlehnung an Schmidt et al. (2009) erhoben. Positive Überzeugungen zum Einsatz digitaler Medien im Unterricht wurden mit sechs Items auf einer vierstufigen Likert-Skala erhoben (Jung & Carstens, 2015).

Erste Ergebnisse zeigen in allen Vergleichsgruppen signifikante Verbesserungen des selbstberichteten technologisch-pädagogischen Wissens im Verlauf des zweitägigen Workshops. Lehrkräfte, die am Reverse Mentoring-Kurs teilgenommen haben, zeigen deskriptiv höhere Zuwächse als Lehrkräfte, die die traditionelle Fortbildungsveranstaltung besucht haben. Eine Varianzanalyse mit Messwiederholung zeigte jedoch keinen statistisch signifikanten Unterschied zwischen den Bedingungen. Studierende, die mit einer Lehrkraft zusammengearbeitet haben, zeigen deskriptiv vergleichbare Zuwächse wie Studierenden, die an traditionellen Hochschulseminaren teilgenommen haben. Auch hier zeigte eine Varianzanalyse mit Messwiederholung keinen signifikanten Unterschied zwischen den Bedingungen. Positive Überzeugungen zur Nutzung digitaler Medien im Unterricht verbesserten sich im Verlauf der Workshoptage signifikant lediglich in der Gruppe der Studierenden in Hochschulseminaren (KG1), ein statistisch signifikanter Unterschied zu Studierenden der Reverse Mentoring-Bedingung konnte jedoch nicht nachgewiesen werden. Lehrkräfte zeigten in beiden Bedingungen keine signifikanten Zuwächse positiver Überzeugungen.

Die Ergebnisse deuten darauf hin, dass Reverse Mentoring geeignet ist, um das selbstberichtete Wissen zum Unterrichten mit digitalen Medien bei Lehrkräften und Lehramtsstudierenden zu fördern. Obwohl sich kein Vorteil der Reverse Mentoring-Bedingung gegenüber den traditionellen Hochschulseminaren und Fortbildungsveranstaltungen zeigte, liefern die Ergebnisse wichtige Erkenntnisse über den Nutzen von Reverse Mentoring als kostengünstige und effektive Methode zur Förderung digitaler Kompetenzen von Lehrkräften in der Aus- und Weiterbildung.

Utility-Value Interventionen zur Unterstützung der TPACK Wissensintegration im Lehramtsstudium

Iris Backfisch[1], Leon Sibley[1], Andreas Lachner[1], Kenneth Tulku Kirchner[1], Christoff Hische[1], Katharina Scheiter[2]
[1]Eberhard Karls Universität Tübingen, [2]Universität Potsdam

Digitale Medien im Unterricht versprechen das Potential den Unterricht lernförderlicher zu gestalten (Buhl et al., 2021). Jedoch erfordert eine sinnvolle Implementierung digitaler Medien in den Unterricht ein hohes Maß an professionellen Kompetenzen bei Lehrkräften (Fraillon et al., 2019). Bisherige Forschung hat gezeigt, dass diese Kompetenzen sowohl aus Wissen als auch motivationalen Facetten bestehen (Backfisch et al., 2020). Das notwendige Wissen wird meist als ein Amalgam verschiedener Wissenskomponenten von technologischen, pädagogischen und fachlichen Inhalten konzeptualisiert (technological-pedagogical-content-knowledge, TPACK, Mishra & Koehler, 2006). Lehramtsstudierende müssen diese Wissenskomponenten integrieren, um digitale Medien lernförderlich im Unterricht einsetzen zu können. Zudem verdeutlichen bisherige Studien, dass vor allem die Nützlichkeitseinschätzung digitaler Medien für das Unterrichten entscheidend für die Qualität des medienbasierten Unterrichts sind (Backfisch et al., 2021). Gleichzeitig zeigte die allgemeine Motivationspsychologie, dass Nützlichkeitsüberzeugungen bei Schüler*innen und Studierende relativ gut durch Utility-Value Interventionen beeinflussbar sind.

In zwei experimentellen Studien wurde daher der Versuch unternommen die Wissensintegration der TPACK Komponenten mit Hilfe einer Utility-Value Intervention zu fördern. Dazu wurden Mathematiklehramtsstudierende (N-Studie 1 = 43, N-Studie 2 = 115) randomisiert entweder der Interventionsgruppe (mit Intervention) oder der Kontrollgruppe (ohne Intervention) zugeteilt. Nach einem Pretest erhielt die Interventionsgruppe die Utility-Value Intervention. Diese bestand aus sechs Statements verschiedener Akteur*innen im Kontext medienbasierten Unterrichtens zur Nützlichkeit von digitalen Medien im Unterricht (d.h., von Wissenschaftler*innen, Lehrerbildner*innen und Politiker*innen). Die Lehramtsstudierenden sollten sich ein Statement auswählen, das ihrer Meinung nach das wichtigste Argument für das Unterrichten mit digitalen Medien enthält. Im zweiten Schritt sollten sie ihre Wahl reflektieren, indem sie beschreiben, inwiefern dieses Statement ihr zukünftiges Unterrichten beeinflussen wird. Danach erhielten beide Gruppen die gleiche Hypertext Lernumgebung zum Unterrichten mit digitalen Medien in Mathematik, die anhand der verschiedenen Wissenskomponenten des TPACK Modells aufgebaut war und jeweils vier Kapitel zu technologischem, pädagogischem und fachlichem Wissen beinhaltete. In dieser Lernumgebung wurden Logfiles erhoben, um die Lernprozesse als Indikator für die Wissensintegration modellieren zu können. Diese Logfiles wurden anhand von epistemischen Netzwerkanalysen (ENA, Shaffer et al., 2016) ausgewertet, die es erlauben das zeitlich nah beieinander liegende Ereignisse (in diesem Fall das Lesen bestimmter Kapitel der Hypertextlernumgebung) darzustellen und mögliche Gruppenunterschiede zu identifizieren. In einem Posttest wurde zusätzlich die Nützlichkeitseinschätzung und das Wissen (Fakten und Transferwissen) erhoben.

Während die Analysen in beiden Studien keine Gruppenunterschiede hinsichtlich der Nützlichkeitseinschätzung und dem Wissen zeigten, wurde anhand der epistemischen Netzwerkanalysen unterschiedliche Lernprozesse in der Lernumgebung zwischen den Gruppen

identifiziert. Über beide Studien hinweg legt die ENA nahe, dass Mathematiklehramtsstudierende in der Interventionsgruppe eher das Wissen der verschiedenen Kapitel integrierten, indem sie zwischen den einzelnen Kapiteln verschiedener Wissenskomponenten wechselten, während Studierende der Kontrollgruppe eher die Kapitel einer Wissenskomponente nacheinander durchlasen. Die Utility-Value Intervention scheint sich also eher implizit auf das Lernverhalten, nicht aber direkt auf die Kompetenzen der Lehramtsstudenten ausgewirkt zu haben. Somit sind diese Studien ein vielversprechender erster Ansatz, Utility-Value Interventionen im Kontext der Lehrerbildung zu implementieren, jedoch ist es wichtig, die Wirkmechanismen von Utility-Value Interventionen weiter zu untersuchen.

D-17: Fünf Jahre nach der Erfassung der ‚Global Competences' in PISA: Entwicklungsperspektiven für zukünftige Erhebungen

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* T03 R04 D10

Fünf Jahre nach der Erfassung der ‚Global Competences' in PISA: Entwicklungsperspektiven für zukünftige Erhebungen

Chair(s): **Jana Costa** (Leibniz Institut für Bildungsverläufe, Germany), **Kathrin Thums** (Leibniz Institut für Bildungsverläufe), **Alscher Pascal** (TU Dortmund)
*Diskutant*in(nen):* **Annette Scheunpflug** (Otto-Friedrich-Universität Bamberg)

Neben klassischen Kompetenzbereichen wie Lesen, Mathematik oder Naturwissenschaften werden im Rahmen des Programms for International Student Assessment (PISA) regelmäßig verschiedene fächerübergreifende Kompetenzen erhoben. Im Jahr 2018 wurden in diesem Rahmen erstmalig die ‚Global Competences' erfasst, um bürgerschaftliches Engagement, interkulturelle Kompetenz und globalisierungsspezifisches Wissen der 15-jährigen Teilnehmenden abzubilden.

Während die Kompetenzmessungen in den klassischen PISA-Domänen auf einem grundlegend erarbeiteten und erprobten Konsens bezüglich relevanter Inhalte und Schwerpunkte aufbauen, betrat die OECD mit den ‚Global Competences' ein Forschungsfeld, welches sich durch dynamische Diskussionen, vielfältige Perspektiven, normative Setzungen und heterogene Diskurse auszeichnet. Es ist daher nicht verwunderlich, dass die Konzeptionierung in wissenschaftlichen Diskursen kritisch wahrgenommen wird (vgl. Kater-Wettstädt & Niemann, 2019; Sälzer & Roczen, 2018; Auld & Morris, 2019; Ress et al., 2022; Costa, 2022).

Im Rahmen des Symposiums wird die Erfassung der ‚Global Competences' aus verschiedenen Perspektiven näher beleuchtet. Ziel ist dabei neben einer kritisch-reflexiven Auseinandersetzung mit dem bestehenden Messinstrument, vor allem das Aufzeigen von Entwicklungspotentialen und übergreifend Ansätze sowie Vorschläge zur Weiterentwicklung anzuregen und zu diskutieren.

Im *ersten Beitrag* wird die Erhebung der ‚Global Competences' über bibliometrische Netzwerkanalysen im wissenschaftlichen Diskurs verortet, indem die im Framework zitierten Werke hinsichtlich ihrer thematischen Schwerpunktsetzungen analysiert werden. Ziel ist es diskursive Schwerpunkte offenzulegen und dadurch potenzielle Anknüpfungspunkte für weiterführende Entwicklungen zu skizzieren.

Im *zweiten Beitrag* folgt eine kritisch-reflexive Auseinandersetzung mit Normativität der ‚Global Competences'. Dazu wird die Selbstdarstellung der OECD mit Blick auf den Umgang mit Normativität aufgezeigt, um darauf aufbauend Herausforderungen zu diskutieren, die mit der Erhebung eines Konstrukts wie Global Competence im Rahmen einer international vergleichenden Studie wie PISA von einer Wirtschaftsorganisation wie der OECD verbunden sind.

Im *dritten Beitrag* wird ein Messinstrument zur Erfassung von komplexen Systemzusammenhängen vorgestellt und daran anknüpfend eruiert, inwiefern eine entsprechende Messung auch für die Erhebung der ‚Global Competences' relevant sein könnte. Mit Vorstellung eines Messinstruments zur Erfassung von Systemkompetenz wird der Blick damit explizit auf die Berücksichtigung von Komplexität gelenkt, um Entwicklungsperspektiven in der Messung von ‚Global Competences' aufzuzeigen.

Zusammenfassend wird in dem Symposium über verschiedene methodische Zugänge ein Einblick in die theoretische Konzeptionalisierung und Verortung der Erfassung der ‚Global Competences' ermöglicht. Über die Vorstellung eines Messinstruments zur Erfassung von Systemkompetenz wird überdies ein erster Blick auf potenzielle (Weiter-)Entwicklungsmöglichkeiten eröffnet. Der zusätzliche Diskussionsbeitrag aus qualitativ-rekonstruktiver Perspektive (Diskutant:in) reflektiert die vorliegenden Beiträge noch einmal im Horizont aktueller Befunde und verspricht eine anregende Diskussion der Entwicklungspotenziale des Gesamtkonzepts.

Beiträge des Symposiums

Verortung der ‚Global Competence' im (bildungs-) wissenschaftlichen Diskurs

Jana Costa[1], Kathrin Thums[1], Alscher Pascal[2]
[1]Leibniz Institut für Bildungsverläufe, [2]TU Dortmund

Die Frage danach, was zukünftige Generationen für die Gestaltung des Zusammenlebens in einer sozial gerechten und ökologisch überlebensfähigen Weltgesellschaft benötigen, ist eine Fragestellung, die für eine nachhaltige Entwicklung von grundlegender Bedeutung ist. Es ist daher begrüßenswert, dass die OECD mit der Erfassung von ‚Global Competences' einen ersten Versuch wagte, sich dem Themen- und Forschungsfeld anzunähern. Gleichzeitig betrat die OECD damit ein Forschungs-Terrain, welches deutlich weniger auf bereits etablierte internationale Definitionen einer Kompetenzdomäne zurückgreifen kann. Die Diskurse können vielmehr als heterogen beschrieben werden und sind international durch verschiedene Denktraditionen geprägt. Im deutschsprachigen Diskurs werden die mit der Entwicklung hin zur Weltgesellschaft verbundenen Herausforderungen in verschiedenen, jedoch eng verzahnten und aufeinander bezogenen Feldern seit Jahrzehnten verhandelt. Dazu zählen das Feld des Globalen Lernens (z.B. Lang-Wojtasik & Scheunpflug, 2005), der Bildung für Nachhaltige Entwicklung (z.B. de Haan, 2007), der Interkulturellen Pädagogik (z.B. Nieke, 2008), der Umweltbildung (z.B. Bögeholz, 1999), der Global Citizenship Education bzw. Development Education (z.B. Bourn, 2014) und der Friedenspädagogik und Menschenrechtsbildung (z.B. Drinkwater et al., 2020). Gerade weil auch bisher kein einheitliches und anerkanntes Kompetenzmodell vorliegt und die Relationierung verschiedener Ansätze weiterhin im Fluss ist, verwundert es nicht, dass die Erfassung der ‚Global Competences' im Diskurs durchaus kritisch wahrgenommen und kommentiert wird (vgl. Kater-Wettstädt & Niemann, 2019; Sälzer & Roczen, 2018; Auld & Morris, 2019; Ress et al., 2022; Costa, 2022).

Ziel des Beitrages ist es, die theoretischen Grundlagen des OECD-Frameworks (OECD, 2019) zu skizzieren, um eine umfassende und transparente Diskussion der (normativen) Prämissen und der (unterschiedlichen) theoretischen Bezugspunkte zu ermöglichen.

Fragestellungen

In dem Beitrag wird eine empirisch fundierte Offenlegung der zugrundeliegenden theoretischen Bezugslinien anvisiert. Dazu wird gefragt:

1. Wie lässt sich die Konzeptionierung der ‚Global Competences' im bildungswissenschaftlichen Diskurs verorten?

2. Auf welchen theoretischen Grundlagen fußt die Modellierung der ‘Global Competences'?

3. Lassen sich diskursive Knotenpunkte identifizieren?

Methodisches Vorgehen

Über eine bibliometrische (Netzwerk-) Analyse (Schneijderberg et al., 2022) wird auf Basis der zitierten Literaturverweise des theoretischen Rahmenpapiers der Zugang zur theoretischen Fundierung möglich. Im Rahmen einer Co-Occourence Analyse (Binary Counting) über das Programm VOSviewer werden verwendete Begriffe in Titel und Abstracts der Publikationen (N = 103 Zitationen) analysiert. Integriert werden Begriffe, die mindestens zweimal genannt werden (N = 400), wobei rein formale Begriffe (z.B. Publikationsjahr oder Füllwörter) ausgeschlossen werden.

Ergebnisse

In den deskriptiven Analysen wird einerseits ein deutlicher Bezug zu anderen Berichten größerer Organisationen (wie UNESCO) sichtbar. Andererseits ist auffällig, dass die überwiegende Anzahl der Publikationen nur einmal im Rahmenpapier zitiert werden. In den bibliometrischen Netzwerkanalyen wird deutlich, dass bestimmte Diskurse repräsentiert werden, aber andere Diskurse unterrepräsentiert bleiben. Neben einem Schwerpunkt auf methodischen Fragen können zwei weitere inhaltliche Diskursfelder identifiziert werden: In den Zitationen werden (1) Fragen kultureller Identität und Interkulturalität fokussiert. In der Netzwerkdarstellung zeigt sich außerdem (2) ein Fokus auf Citizenship (education). Beide inhaltlichen Diskursfelder sind über die UNESCO miteinander verbunden.

Diskussion

Grundsätzlich geht mit jeder standardisierten Erhebung notwendigerweise eine bestimmte theoretisch-konzeptionelle Setzung und die Reproduktion spezifischer Perspektiven einher. Es lässt sich in einer empirisch-quantitativen Studie nicht vermeiden und ist für die Operationalisierung essenziell, einen theoretischen Referenzrahmen aufzuspannen. Im Kontext der skizzierten diskursiven Schwerpunkte von ‚Global Competences' lässt sich allerdings eine Diskursverengung bzw. eine Komplexitätsreduktion erkennen, welche angesichts der vielfältigen Diskurse und der komplexen Lernherausforderungen in der globalisierten Weltgesellschaft zum Nachdenken über (Weiter-) Entwicklungspotenziale einlädt. Der klare Fokus auf Aspekte interkultureller Kompetenz und die weniger intensive Auseinandersetzung mit Fragen von globaler Gerechtigkeit, Machtverhältnissen und nachhaltigkeitsbezogenen Fragestellungen lässt ein spezifisches Verständnis von Globalität sichtbar werden, welches in zukünftigen Erhebungen noch einmal theoretisch geschärft werden muss. Die skizzierten Diskursfelder erlauben in diesem Zusammenhang Anregungen zur Weiterentwicklung zu formulieren und zu diskutieren.

Die Normativität der „Global competences" – die Notwendigkeit kritisch-reflexiver Ansätze

Lydia Kater-Wettstädt
Uni Potsdam

Die Frage der expliziten normativen Ausrichtung von Global Competence bzw. verwandter Konzepte begleitet den Diskurs z.B. um eine Bildung für nachhaltige Entwicklung (BNE) oder Globales Lernen (GL) von Beginn an. Verbunden damit ist die Frage wie weit der Anspruch, auf nachhaltiges bzw.

bestimmten Gerechtigkeitsansprüchen entsprechendes Handeln durch Bildung hinzuwirken, gehen darf.

Im Rahmen einer BNE wird z.B. auf der einen Seite Bildung für eine nachhaltige Entwicklung wortwörtlich als Handlungsrichtungen vorgebend verstanden. Auf der anderen Seite steht eine Bildung als nachhaltige Entwicklung (Scott & Gough, 2003; Vare & Scott, 2007), die vor allem auf kritisches Denken abzielt, das auch auf das Aufdecken von Widersprüchen zielt (z.B. in Bezug auf Entwicklungsparadigmen). Aufgrund des „moralischen Imperativs“ (Rost, 2002) ist gerade im Kontext BNE oder GL der Vorwurf einer Indoktrination oder Instrumentalisierung oft nicht weit (Jickling, 2003; van Poeck & Vandenabeele, 2014).

Die Messung von Global Competence stellt im Vergleich zu bisherigen Konstrukten diesbezüglich eine Neuerung dar: Auch wenn zuvor bereits Einstellungen, Verhalten, Motivation und Werte erhoben wurden, wurden diese bisher hauptsächlich zur Erklärung kognitiver Unterschiede genutzt. Jetzt aber beschreibt die OECD in der Rahmenkonzeption, dass jede einzelne der vier Komponenten von Global Competence Einstellungen (z.B. Offenheit und eine Haltung des Respekts für kulturelle Unterschiede) und Werte (Wertschätzung von Menschenwürde und kultureller Vielfalt) umfasst. Gleichzeitig verweist die OECD darauf, diese Komponenten nicht messen zu können (OECD 2018). Dies führt dazu, dass sich für jede der Komponenten essenzielle Teilaspekte der Messung entziehen. Dies war bisher in keiner der bisherigen PISA Rahmenkonzeptionen der Fall (Roczen & Kater-Wettstädt, 2022).

Fragestellung

Angesichts der spezifischen kompetitiven Ausrichtung auf die Verwertbarkeit von Bildung im Sinne ökonomischer Prosperität (z.B. Grotlüschen, 2018, Kater-Wettstädt &Niemann, 2019) stellen sich die Fragen wie mit Werten und der Normativität im Kontext der Erfassung von Global Competences umgegangen wird und wie diese reflektiert und diskutiert werden.

Methodisches Vorgehen

In einer systematischen Dokumentenanalyse (Wolff, 2013) der Rahmenpapiere 2016/2018 und Ergebnisdarstellung werden die zentralen Argumentationsfiguren, ihre Konsistenz oder auch Widersprüchlichkeit rund um das Thema Werte und Normativität herausgearbeitet und aufgezeigt, wie sich die OECD versucht in diesem Spannungsfeld zu positionieren.

Ergebnisse

Bereits in der Begründung für die Erhebung des Konzeptes findet sich gleich hinter dem Anspruch eines harmonischen Zusammenlebens die ökonomische Kontextuierung von Bildung in der Figur des Bestehens auf dem sich wandelnden Arbeitsmarkt. Im weiteren Verlauf wird als Werte-Grundlage die Menschenwürde und die Allgemeine Erklärung der Menschenrechte gesetzt. Die OECD räumt einer entsprechenden Kritik vorwegnehmend den westlichen Ursprung der Dokumente ein. Zugleich findet aber keinerlei Verhältnisbestimmung zur eigenen Perspektivität statt, außer eine Ergänzung um weltlich-religiöse Legitimationsbezüge, z.B. in Bezug auf Ubuntu (OECD, 2018, S. 19).

Im Kern der Argumentation wird auf normativer Ebene Pluralität von Meinungen und eine notwendige Offenheit propagiert. Zu keinem Zeitpunkt wird allerdings die eigene Standortgebundenheit kritisch

reflektiert und sich mit der Gefahr der Instrumentalisierung oder Indoktrination auseinandergesetzt, z.B. vor dem Hintergrund des Spannungsverhältnisses zwischen dem eigenen Entwicklungsverständnis der OECD und einer nachhaltigen, gerechten Entwicklung.

Werte und damit z.B. auch Normalitätsvorstellungen müssen allerdings insbesondere in pädagogischen Kontexten transparent gemacht und reflektiert werden (Lahsen &Turnhout, 2021). Die vorgenommenen Setzungen der OECD aber werden in ihren unterschiedlichen Spannungsverhältnissen allerdings nicht bearbeitet.

Für zukünftige Erhebungen, im Kontext von PISA und darüber hinaus, wäre es von Bedeutung die eigenen Normen und Wertevorstellungen kritisch zu hinterfragen und deren Konsequenzen für das Erhebungsdesign auszuloten., z.B. im Hinblick auf die Anlage von PISA als internationaler Leistungsvergleich (z.B. zur Problematik der (Nicht-)Beteilung von Ländern, Lange, 2015), die ebenfalls in ihrer kompetitiven (und exklusiven) Form nicht hinterfragt wird. Über die Reflexion hinaus bedarf es zudem sicherlich einer Weiterentwicklung der Instrumente im Sinne eines emanzipatorischen Bildungsverständnisses, das z.B. nach erstrebenswerter und zukunftsfähiger Entwicklung fragt.

Systemkompetenz als zentrales Element von Globaler Kompetenz?

Nina Roczen[1], Johannes Hartig[1], Rainer Mehren[2], Janis Fögele[3], Frank Fischer[4]
[1]Leibniz-Institut für Bildungsforschung und Bildungsinformation, [2]Universität Münster, [3]Universität Hildesheim, [4]Justus-Liebig-Universität Gießen

Theoretischer Hintergrund

Der aktuelle Club of Rome Bericht (Dixson-Declève et al., 2022) betont als Aufgabe von Bildung die Förderung des Verständnisses für komplexe Systeme, was auch in verschiedenen Ansätzen in der Literatur zu Bildung für Nachhaltige Entwicklung (BNE) als zentrale Schlüsselkompetenz für den Umgang mit globalen Herausforderungen gilt (Wiek et al., 2011; UNESCO, 2017; Riekmann, 2018). Auch in der im Rahmen der PISA Studie durchgeführten Messung von „Global Competence“ (OECD, 2018) wurde der Umgang mit Komplexität thematisiert mit Kompetenzen für eine „vernetzte, diverse und sich schnell verändernden Welt“ – Allerdings wurde die Kompetenz zum Umgang mit komplexen Systemen nicht explizit gemessen. In diesem Beitrag wird ein Messinstrument für Systemkompetenz vorgestellt das es erlaubt, Systemkompetenz einem Large Scale-Setting wie z.B. dem PISA Global Competence-Assessment zu erheben. Bisherige Instrumente zur Messung von Systemkompetenz wären für ein solches Setting zu aufwendig in der Auswertung (Roczen et al., 2021).

Systemkompetenz wird als die Fähigkeit verstanden, komplexe Systeme zu analysieren, zu verstehen, vorherzusagen sowie Maßnahmen zur Regulation zu erkennen (Mehren et al., 2016). Wir haben Testaufgaben zu den Themenbereichen Elektroschrott, Klimawandel und Korallensterben, Megastädte, Fleischkonsum und Regenwald, Textilproduktion und Übertourismus entwickelt. Diese verwenden ausschließlich geschlossene Antwortformate wie Multiple-Choice-Aufgaben oder Aufgaben mit Pfeildiagrammen (Concept Maps) zur Darstellung von Beziehungen innerhalb von Systemen. Basierend auf den Ergebnissen einer qualitativen Pilotierung in Cognitive Laboratories und durch das Feedback von Expert*innen wurde der Test nochmals angepasst, bevor wir einen Pilottest zur psychometrischen Überprüfung des Tests durchgeführt haben und das Instrument nochmals

entsprechend überarbeitet haben. Anschließend haben wir nochmals Daten von N=1000 15-jährigen Schülerinnen und Schülern erhoben, um die folgenden Forschungsfragen beantworten zu können:

(1) Kann der Test reliabel Systemkompetenz messen?

(2) Decken die Schwierigkeiten der Testaufgaben das Kompetenzspektrum von 15-Jährigen ab?

(3) Ist es möglich, Systemkompetenz mit nur einer Dimension zu erfassen?

(4) Lassen sich die empirischen Schwierigkeiten mit den bei der Aufgabenentwicklung intendierten Kompetenzniveaus vorhersagen?

Methode

Wir haben Daten von N = 1098 Schülerinnen und Schülern der 9. Klasse aller Schularten erhoben. Dabei wurde ein Testheftdesign verwendet, bei dem die Schüler*innen jeweils nur 3 der insgesamt 6 Testaufgaben bearbeiten mussten. Wir prüfen die Konformität der Daten mit dem eindimensionalen (Partial-Credit) Rasch-Modell, bestimmen die psychometrischen Eigenschaften der Items, die Reliabilität der Kompetenzschätzungen und die Passung der Verteilungen der Itemschwierigkeiten und der Kompetenzen der Schülerinnen und Schüler. Darüber hinaus prüfen wir die Dimensionalität des Instruments mit einer konfirmatorischen Faktorenanalyse (CFA) und untersuchen den Zusammenhang zwischen intendierten Kompetenzniveaus und empirischen Aufgabenschwierigkeiten regressionsanalytisch.

Ergebnisse und ihre Bedeutsamkeit

Der Test erreicht mit drei Aufgaben pro Schüler*in eine EAP-Reliabilität von 0.90. Die Itemkennwerte sprechen für eine gute Passung zum Partial Credit-Modell, und die Ergebnisse der CFA stützen die Annahme einer eidimensionalen Skala. Durch die bei der Konstruktion intendierten Kompetenzniveaus der Aufgaben lassen sich mehr als 20% der Varianz in den empirischen Itemschwierigkeiten erklären.

Die Ergebnisse werden diskutiert im Hinblick darauf, was eine auf diese Weise erfasste Systemkompetenz leisten könnte im Sinne einer Weiterentwicklung des Global Competence Assessments. Außerdem soll diskutiert werden welche Desiderate noch bleiben, z.B. im Hinblick auf die Abbildung von Unsicherheit oder von nicht-kognitiven Aspekten des Umgangs mit Komplexität.

D-18: Evaluation digitaler Lernumgebungen zur Förderung der Klassenführungskompetenz bei Lehramtsstudierenden

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* R11 T00 D01

Evaluation digitaler Lernumgebungen zur Förderung der Klassenführungskompetenz bei Lehramtsstudierenden

Chair(s): **Robin Junker** (Westfälische Wilhelms-Universität Münster)

Digitale Formate haben in unterschiedlichen Bereichen unseres gesellschaftlichen Zusammenlebens Eingang gefunden (Boellstorff, 2015). Auch im Bildungssystem werden digitale Lehr-Lern-Formate zunehmend eingesetzt, um spezifischen Anforderungen gerecht zu werden. Dazu gehören zum Beispiel der steigende Bedarf nach Individualisierung des Lernverlaufs (Landenfeld et al., 2019) oder Flexibilität beim Lernen (Lester & King, 2009). Auch im Rahmen der universitären Lehrkräftebildung werden die Potentiale digitaler Lehr-Lern-Formate erkannt und beantwortet. Durch deren Einsatz kann unter anderem des Erlernen zentraler Qualitätsmerkmale guten Unterrichts unterstützt werden. Ein solches essentielles Merkmal stellt die Klassenführung dar (Emmer und Stough 2001; Klieme et al. 2001). Sie hängt direkt mit der Lernmotivation (Kunter et al. 2007; Helmke 2007), der Leistung (Wang et al. 1990) und dem Kompetenzzuwachs (Baumert et al. 2003; Hattie 2012) der Lernenden zusammen. Um Wissen, Wahrnehmungs- und Handlungskompetenzen zum Thema Klassenführung zu fördern, werden zunehmend digitale Lehr-Lern-Umgebungen eingesetzt, die in den einzelnen Beiträgen vorgestellt und evaluiert werden.

Beitrag 1 von Lehmkuhl et al. beispielweise stellt dar, wie die mehrperspektivische professionelle Wahrnehmung von Klassenführung und Lernunterstützung im naturwissenschaftlichen Sachunterricht mit Hilfe eines digitalen, videobasierten Selbstlernmoduls inklusive spezifischer Unterstützungstools gefördert werden kann. In Beitrag 2 wird hingegen von Huang et al. dargestellt, inwiefern Videos bzw. virtual reality eingesetzt werden können, um das Interesse sowie die Selbstwirksamkeit in Bezug auf Klassenführung zu steigern. Daran anschließend wird in Beitrag 3 von Böhnke et al. untersucht, ob Reasoning-Kompetenzen von Lehramtsstudierenden gefördert werden können, wenn störungskritische Ereignisse in Form von Textvignetten in einer digitale Lernumgebung präsentiert und analysiert werden sollen. In Beitrag 4 beschreiben Salzmann et al. den Einfluss der Verwendung eines selbst entwickelten digitalen Selbstlernmoduls zum Thema Klassenführung mit inszenierten Unterrichtsseqzuenzen auf die Klassenführungsexpertise und die Selbstwirksamkeitsüberzeugung von Lehramtsstudierenden. Zuletzt soll in Beitrag 5 von Junker et al. erläutert werden, wie eine so komplexe Aufgabe wie das Lernen der professionellen Wahrnehmung von Klassenführung in digitalen Lernumgebungen mit Video Modeling und Prompts unterstützt werden kann sowie welche Auswirkungen diese Unterstützung auf die kognitive Belastung und die Motivation der Lernenden hat.

Alle Beiträge liefern wichtige Erkenntnisse zur Förderbarkeit von klassenführungsbezogenen Kompetenzen in digitalen Lehr-Lern-Formaten. Es soll diskutiert werden, welche Prinzipien auf der Basis dieser empirischen Ergebnisse für die Gestaltung digitaler Lernumgebungen zur Förderung von Klassenführungskompetenzen bei Lehramtsstudierenden abgeleitet werden können sowie welche weiteren Forschungsschwerpunkte man anhand der gewonnenen Erkenntnisse identifizieren kann.

Förderung der mehrperspektivischen Wahrnehmung von Klassenführung und Lernunterstützung mithilfe digitaler, videobasierter Lernmodule – Zum Effekt spezifischer Unterstützungstools

Alena Lehmkuhl, Jennifer Janeczko, Verena Zucker, Robin Junker, Manfred Holodynski, Nicola Meschede
Westfälische Wilhelms-Universität Münster

Verschiedene Qualitätsdimensionen bedingen einander wechselseitig im Unterricht (Kunter & Voss, 2011) – dies trifft auch auf den naturwissenschaftlichen Sachunterricht zu (Fauth et al., 2014). So stellt z.B. die Klassenführung einer Lehrkraft eine notwendige, wenn auch keine hinreichende Bedingung für eine wirksame Lernunterstützung dar (Klieme et al., 2001). Eine entsprechend mehrperspektivische professionelle Unterrichtswahrnehmung von Wechselwirkungen kann dabei in Anlehnung an Blömeke et al. (2015) sowie Seidel und Stürmer (2014) als bedeutsam für das eigene Unterrichtshandeln erachtet werden. Ausgehend von Junker et al. (2020) ist damit die Fähigkeit gemeint, diese Wechselwirkungen im komplexen Unterrichtsgeschehen erkennen und anschließend (1) beschreiben, (2) interpretieren, (3) bewerten sowie (4) passende Handlungsalternativen generieren zu können. Bisherige Studienergebnisse verdeutlichen, dass eine Förderung der professionellen Unterrichtswahrnehmung mithilfe der Analyse von Unterrichtsvideos bereits im Lehramtsstudium möglich ist (z.B. Gold et al., 2020; Sunder et al., 2016), jedoch aufgrund unterschiedlicher Entwicklungsverläufe der Studierenden (Stürmer et al., 2016) Unterstützungsangebote z.B. im Sinne des Cognitive-Apprenticeship-Ansatzes (Collins et al., 1989; Junker et al., 2020) bedarf. Eine digitale Umsetzung solcher Unterstützungstools könnte einen Mehrwert zur Kompetenzförderung bieten, da diese neue Möglichkeiten zur adaptiven Bereitstellung individueller Hilfestellungen (z.B. in Form von Modellierungsvideos oder Prompts) eröffnet, die den Lernprozess begünstigen können (García-Cabrero et al., 2018). Bislang wurde der Einsatz solcher Tools in diesem Förderkontext noch nicht untersucht.

Im Projekt ProdiviS (https://prodivis.de/) wurde vor diesem Hintergrund eine Förderung der mehrperspektivischen professionellen Wahrnehmung von Klassenführung und Lernunterstützung im naturwissenschaftlichen Sachunterricht mithilfe digitaler, videobasierter Lernmodule angestrebt. Der Fokus der Untersuchung lag auf der Integration zusätzlicher Unterstützungstools bei der Analyse videografierten Unterrichts mit besonderem Fokus auf Modellierungsvideos sowie spezifischen Prompts. Insgesamt wurden für die Untersuchung fünf Lernmodule konzipiert, die eine komplexitätssteigernde Förderung der mehrperspektivischen professionellen Unterrichtswahrnehmung ermöglichen sollen. Die vorliegende Studie untersucht das Förderpotenzial der zusätzlich integrierten Unterstützungstools, die in diese Lernmodule eingebettet wurden.

Im Rahmen der laufenden quasi-experimentellen Interventionsstudie mit Prä-Post-Follow-Up-Design besuchten 103 Masterlehramtsstudierende des Sachunterrichts an Grundschulen eine Lehrveranstaltung an der Westfälischen Wilhelms-Universität Münster (SoSe 2021 bis SoSe 2022), in der sie die fünf digitalen Lernmodule bearbeiteten. Ein Teil der Studierenden (n = 53) bearbeitete die Module mit spezifischen Unterstützungstools bei der Videoanalyse, ein Teil der Studierenden durchlief die Module ohne ergänzende Unterstützungstools (n = 50). Zusätzlich wurde eine unbehandelte Kontrollgruppe (N = 43) rekrutiert, die ein reguläres universitäres Semester an anderen Universitäten in Nordrhein-Westfalen absolvierte.

Die mehrperspektivische professionelle Wahrnehmung von Klassenführung und Lernunterstützung wird anhand einer strukturierten, offenen Videoanalyse erfasst. Die Auswertung der offenen Antworten der Teilnehmenden erfolgt auf Grundlage der qualitativen Inhaltsanalyse nach Mayring (2015). Hierfür wurden ein umfangreiches Kodiermanual sowie ein Expert*innenrating entwickelt, anhand derer drei Kodiererinnen die Antworten der Studierenden einschätzen. Die Interrater-Reliabilität (22% Doppelkodierung) liegt in einem guten bis sehr guten Bereich (Beschreiben: Fleiss' κ = .681, Interpretieren: ICC = .864, Bewerten: ICC = .866, Handlungsalternativen: ICC = .911). Die letzte Follow-Up-Erhebung der Kohorte SoSe 2022 erfolgt Mitte Oktober 2022, sodass die Kodierung der offenen Antworten im November abgeschlossen werden kann.

Zusätzliche subjektive Kompetenzeinschätzungen der Studierenden aus dem Sommersemester 2022 deuten bereits auf einen Kompetenzzuwachs durch den Einsatz der Lernmodule mit zusätzlichen Unterstützungstools hin (Zucker et al., in Vorb.). Die Ergebnisse der offenen Analysen bzgl. der Förderung der mehrperspektivischen Unterrichtswahrnehmung in Bezug auf den Einsatz digitaler Unterstützungstools bei der Analyse von Unterrichtsvideos werden im Vortrag vorgestellt.

Comparing video and virtual reality as tools for fostering interest and self-efficacy in classroom management: Results of a pre-registered experiment

Yizhen Huang[1], Eric Richter[1], Thilo Kleickmann[2], Dirk Richter[1]
[1]Universität Potsdam, [2]Universität Kiel

Theoretischer Hintergrund

Video is a widely used medium in teacher training for situating student teachers in classroom scenarios (Gaudin & Chaliès, 2015). Although the emerging technology of virtual reality (VR) provides similar, and arguably more powerful, capabilities for immersing teachers in lifelike situations, its benefits and risks relative to video formats have received little attention in the research to date. Among the many potential outcomes of video and VR-based training, we focused on student teachers' interest and self-efficacy in classroom management as the main variables of interest in the present study, given their theoretical significance and attested value for teachers' professional development (for a review, see Morris et al., 2017).

Being interested in something is the starting point and also the driving force of learning (Palmer, 2004). Interest is a motivational variable describing the psychological state of (re-)engaging with particular topics (Hidi & Renninger, 2006). Individuals' beliefs about their competence, that is, self-efficacy, form another key motivational construct in learning (Bardach & Klassen, 2021; Eccles & Wigfield, 2002). Teachers' self-efficacy can be defined as beliefs about their ability to "organize and execute the courses of actions" (Bandura, 1997, p. 3) necessary to be effective in instructional situations.

Fragestellung

Our study aimed to examine how changes in situational interest and self-efficacy in classroom management differ when student teachers undertake an identical classroom management task in a video versus VR environment. We were also interested in the perceived extraneous cognitive strain associated with the medium.

Methode

Forty-nine student teachers (24 female, 23 male, 2 missing) aged 18–43 years (M = 22.4, SD = 4.2) were recruited from the teacher preparation program at a public German university. Both the procedure and the classroom management task were identical across groups. The general procedure for both groups took three steps. Participants first completed a 20-minute pretest questionnaire about their demographics, situational interest, and self-efficacy in classroom management. Second, participants performed a 10-minute classroom management task during which they needed to respond to typical classroom disruptions presented either with video snippets or in a VR classroom (Figure 1). Finally, after the classroom management task, all participants completed a posttest questionnaire that included the pretest questions and also cognitive load ratings of the medium. We assessed triggered interest in classroom management with a 5-item measure from Rakoczy et al. (2020). Classroom management efficacy was assessed with a five-item measure from Pfitzner-Eden et al. (2014).

Ergebnisse

Linear mixed (effects) modeling was used to evaluate the main effect of condition (VR vs. video) and time (pretest vs. posttest) on triggered interest and self-efficacy. We found that both the condition and time significantly predicted the triggered interest meaning that both groups had significant increments from pretest to posttest. A significant interaction ((β̂ = −0.51, t(45) = −2.90, p = .006, d = −0,97) between group and time for the triggered interest demonstrated that the increment in triggered interest in the VR group was significantly larger than the video group. This finding was similar for self-efficacy ((β̂ = −0.36, t(45) = −2.56, p = .014, d = -0,.75). The VR group (EMM = 1.74 , SE = .07) also reported higher extraneous cognitive load than the video group (EMM = 1.47, SE = .07), F(1, 87) = 8.28, p = .005, η^2 = .08 (medium to large effect size).

Bedeutung

The findings demonstrated that immersive VR could be an authentic teacher training environment that reflects the way professional competencies are applied in a real-life classroom. Concurrently, we also found that the VR simulation invoked higher extraneous cognitive load than video, indicating that the feeling of “being there” may be a double-edged sword for learning.

Unterrichtsstörungen richtig interpretieren. Eine experimentelle Studie zur Förderung von Reasoning-Kompetenzen bei Lehramtsstudierenden

Anja Böhnke, Maxie Kilbury, Felicitas Thiel
Freie Universität Berlin

Störungen im Unterricht stellen vor allem unerfahrene Lehrkräfte vor große Herausforderungen. Für einen funktionalen Umgang mit Unterrichtsstörungen ist eine gute professionelle Unterrichtswahrnehmung entscheidend (Barth, 2017). Neben dem Identifizieren von kritischen Ereignissen (Noticing), ist vor allem auch deren adäquate und theoriebasierte Interpretation (Reasoning) zentraler Aspekt der professionellen Wahrnehmung (Sherin, 2001). Bisherige Studien zeigen, dass Noviz:innen im Gegensatz zu Expert:innen große Schwierigkeiten haben, vom sichtbaren Verhalten Schlüsse auf dahinter liegende Motive zu ziehen. Ihre Problemanalysen fallen oft unvollständig oder fehlerhaft aus und dementsprechend häufig reagieren sie in Störungssituationen dysfunktional (Stahnke & Blömeke, 2021; Thiel et al., 2012).

Um Reasoning-Kompetenzen von Lehramtsstudierenden zu fördern, wurden als Fallbeispiele Unterrichtsszenarien mit störungskritischen Ereignissen in Form von Textvignetten entwickelt und in eine digitale Lernumgebung implementiert. Die Textvignetten entstanden auf der Basis von theoretischen Modellen zum Klassenmanagement (Ophardt & Thiel, 2013) und zu Unterrichtsstörungen (Böhnke, Bach & Thiel, in Vorb.). Entscheidend ist hier die Unterscheidung verschiedener Typen von Störungen und entsprechender Strategien zur Störungsintervention (Thiel, 2016), dabei werden Störungen danach klassifiziert, welche Motive von Schüler:innen einem störungskritischen Verhalten zugrunde liegen. Für jede Textvignette wurde für die digitale Lernumgebung ein Videotutorial (protagonistisches Tutorial, Valtentin, 2021) produziert, in dem eine Expertin die Unterrichtsszenarien analysiert. In diesen Tutorials werden theoretische Kernkonzepte aufgegriffen und ihre Relevanz in Bezug auf die dargestellten Unterrichtsverläufe verdeutlicht. So analysiert die Expertin einerseits, welche Hinweise auf die Motive von störungskritischem Schüler:innenverhalten in den Textvignetten zu erkennen sind und andererseits, wie die Handlungsstrategien der Lehrkraft zur Störungsintervention zu bewerten sind und welche Handlungsstrategien in den konkreten Situationen gegebenenfalls funktionaler sein könnten. In diesem Sinne werden in den Tutorials Reasoning-Prozesse modelliert.

Anhand einer Stichprobe von ca. 150 Lehramtsstudierenden, die die digitale Lernumgebung bearbeiteten, wurde ein experimentelles Prä-post-Design realisiert. Es sollte geprüft werden, ob das zusätzliche Betrachten der Expert:innen-Tutorials die Reasoning-Kompetenzen der Studierenden fördern kann. Dazu erhielten alle Teilnehmer:innen eine theoretische Einführung in das Thema Klassenmanagement und Umgang mit Unterrichtstörungen in der digitalen Lernumgebung. Anschließend bearbeiteten sie die Textvignetten (N = 7), in denen kritische Unterrichtsereignisse theoriebasiert interpretiert werden sollten. Eine der Vignetten wurde sowohl zu Beginn als auch zum Ende bearbeitet (Prä-post-Vignette). Die Experimentalgruppe erhielt zusätzlich zu jeder Textvignette das entsprechende Expert:innen-Tutorial (nicht für die Prä-post-Vignette). Die Antworten der Studierenden wurden von geschulten Ratern nach einem Manual kodiert. Um zusätzlich das Wissen zu Klassenmanagement und dem Umgang mit Störungen zu erfassen, wurde ein standardisierter Wissenstest eingesetzt (Böhnke, Bach & Thiel, in Vorb.).

Die Ergebnisse einer einfaktoriellen MANOVA zeigten einen signifikanten Unterschied zwischen der Experimental- und der Kontrollgruppe ($F(2,45) = 4.66$, $p = .014$, partielles $\eta^2 = .172$, Pillai trace = .17). Die Analyse der univariaten Effekte zeigte, dass sich die Experimentalgruppe bei der Bearbeitung der Prä-post-Vignette signifikant verbesserte, während dies für die Kontrollgruppe nicht der Fall war. In Bezug auf den Wissenstest verbesserten sich beide Gruppen, die Experimentalgruppe aber signifikant stärker. Es sollen anhand der anderen bearbeiteten Vignetten Wachstumskurvenmodelle für Experimental- und Kontrollgruppe gerechnet werden. Die Ergebnisse werden diskutiert.

Beschreibung und Evaluation eines videobasierten Selbstlernmoduls für Lehramtsstudierende zur Schulung der professionellen Unterrichtswahrnehmung mit dem Fokus Klassenführung

Sara Salzmann, Sarah Strauß, Johannes König, Kai Kaspar
Universität zu Köln

Theoretischer Hintergrund

Klassenführung ist eines der zentralen Qualitätsmerkmale für guten Unterricht (Hattie, 2009), sowohl bezogen auf die Leistungen, Lernzufriedenheit und Motivation der Schüler:innen (z.B. Wang et al., 1990; Aldrup et al., 2018) als auch für das Wohlbefinden der Lehrperson (z.B. Barth, 2017; Aloe et al., 2014). Die Vorbereitung von Lehramtsstudierenden auf diesen zentralen Aufgabenbereich kann in der universitären Ausbildung als noch nicht adäquat zusammengefasst werden, Studierende selbst fühlen sich nur unzureichend auf diese Anforderung vorbereitet (z.B. Poznanski et al., 2018). Um das komplexe Zusammenspiel verschiedener klassenführungsrelevanter Bedingungen realistisch darstellen und bereits in der Lehramtsausbildung nutzen zu können, wurde im Rahmen des Verbundprojekte ProdiviS - Förderung der professionellen Wahrnehmung in digitalen, videobasierten Selbstlernmodulen am Standort Köln ein digitales Selbstlernmodul für Studierende entwickelt. Durch die Arbeit mit animierten Unterrichtsvideos soll fallbasiertes Lernen, mit dem Ziel der Förderung der Professionellen Wahrnehmung im Hinblick auf Klassenführung in drei Schlüsselmomenten der Klassenführung – Übergängen, Zielklarheit der Instruktion sowie Regeln und Routinen - ermöglicht werden. Für das Lernen in der digitalen Umgebung wurden inszenierte Unterrichtssituationen verwendet, da diese, im Gegensatz zu authentischen Unterrichtssituationen, sowohl eine gezielte und prägnante als auch eine theoretisch fundierte, holistische Darstellung von klassenführungsrelevanten Strategien ermöglichen (Bönte et al., 2019). Ergebnisse aus anderen Studien haben gezeigt, dass sowohl durch die Arbeit mit Unterrichtsskripten als auch mit Unterrichtsvideos für das Themenfeld Klassenführung Lernzuwächse bei Lehramtsstudierenden festgestellt werden konnten (Syring et al., 2015; Gold et al., 2016; Kramer et al., 2017, 2020). Verschiedene Studien konnten zeigen, dass der Einsatz von Unterrichtsvideos in der Lehrer:innenausbildung unter anderem einen positiven Einfluss auf Selbstwirksamkeitsüberzeugungen und die Professionellen Wahrnehmung von Klassenführung hat (Gold et al., 2017) und im Vergleich zu textbasierten Unterrichtsskripten bei den Lernenden die kognitive Aktivierung und die Freude an der Teilnahme fördern (Kramer et al., 2017). Zudem fehlen in der Lehrkräfteausbildung digitale Lernumgebungen, die bspw. in der Praxisphase den geringen Präsenzzeiten der Studierenden, oder dem Bedarf an flexiblen, ortsunabhängig nutzbaren und digitalen Lernumgebungen Rechnung tragen. Von insgesamt 30 animierten Unterrichtsvideos wurden im Sommersemester 2022 erstmalig 15 Unterrichtsvideos in Form eines digitalen Selbstlernmoduls in der Lehre eingesetzt und umfangreich evaluiert.

Fragestellungen

Im Fokus der Untersuchung steht die Begleitevaluation des Selbstlernmoduls.

F1: Wie beurteilen die Studierenden die Arbeit mit dem Selbstlernmodul im Hinblick auf die kognitive Belastung, User Experience und intrinsische Motivation?

F2: Verbessert sich das Wissen über Klassenführungsexpertise sowie die Selbstwirksamkeits-überzeugungen über Klassenführung? In welchem Zusammenhang stehen die Ergebnisse mit Lerngelegenheiten der bisherigen Lehramtsausbildung?

F3: In welchem Zusammenhang stehen die Arbeit mit dem digitalen Selbstlernmodul und die Einstellung zum digitalen Lernen und der Lernzufriedenheit?

Stichprobe

Insgesamt sieben Master-Seminare der Universität zu Köln (vier Interventions-, drei Kontrollgruppen) haben an einer ersten Begleitevaluation des digitalen Selbstlernmoduls im Sommersemester 2022 zur Förderung der professionellen Wahrnehmung von Klassenführung, im Rahmen einer Prä- und Postmessung sowie einer Usability-Befragung teilgenommen (Längsschnitt: n = 16 Interventionsgruppe, n = 24 Kontrollgruppe). Zusätzlich haben Studierende der Interventionsgruppe nach jeder Intervention an einer kurzen Befragung zur Usability teilgenommen.

Methode und Instrumente

Eingesetzt wurden sowohl bereits etablierte Instrumente, wie das videobasierte CME-Instrument (Classrooom Management Expertise, König, 2015), Selbstwirksamkeitsüberzeugungen über Klassenführung (Pfitzner-Eden et al., 2014), Intrinsische Motivation (Wilde et al., 2009), User Experience (Hassenzahl et al., 2003), Intrinsische kognitive Belastung (Kramer et al., 2020), Extrinsische kognitive Belastung (Klepsch, Schmitz & Seufert, 2017) als auch im Rahmen des Projektes ProdiviS neu entwickelte oder adaptierte Instrumente, wie die Einstellung zum digitalen Lernen, die Lernzufriedenheit (in Anlehnung an Nagy, 2018) oder Lerngelegenheiten. Im Rahmen von Gruppen- und Prä-/Postvergleichen (Varianzanalysen und Regressionsmodellen) werden die oben genannten Fragestellungen überprüft.

Ergebnisse

Erste Ergebnisse der derzeit noch nicht abgeschlossenen Datenanalyse zeigen für die verwendeten Skalen zufriedenstellende Reliabilitäten. Zudem zeigen sie für die Wahrgenommene Wahlfreiheit (Unterskala der Intrinsischen Motivation) und die Lernzufriedenheit erwartungskonform signifikante Unterschiede im Zuwachs zwischen der Kontroll- und Interventionsgruppe. Für die Selbstwirksamkeitsüberzeugungen über Klassenführung, die Einstellungen zum digitalen Lernen und das Interesse/Vergnügen (Unterskala der Intrinsischen Motivation) zeigen sich sowohl in der Interventions- als auch in der Kontrollgruppe signifikante Zuwächse zwischen beiden Messzeitpunkten.

Ausblick

Die Ergebnisse sind insbesondere relevant für die (Weiter-) Entwicklung weiterer videogestützter Tools und digital gestütztes Selbstlernen im Lehramtsstudium.

Zum Zusammenhang von Video-Modeling und Prompts mit der Motivation und der kognitiven Belastung beim digitalen Lernen der professionellen Wahrnehmung von Klassenführung

Junker Robin, Jennifer Janeczko, Alena Lehmkuhl, Verena Zucker, Manfred Holodynski, Nicola Meschede
Westfälische Wilhelms-Universität Münster

Die Verwendung von digitaler Technologie ist in unterschiedlichen Gesellschaftsbereichen essentiell geworden (Boellstorff, 2015). Da Bildungseinrichtungen Lernende auf spätere Anforderungen in der Gesellschaft vorbereiten sollen (Aslan & Duruhan, 2021), werden daher in den letzten Dekaden zunehmend digitale Lernumgebungen (VLE) verwendet (Stevens et al., 2021). Es existieren aber auch andere Vorteile solcher Lernumgebungen, wie zum Beispiel die Möglichkeit zu individualisierten Lernpfaden (Landenfeld et al., 2019), asynchroner Bearbeitung (Lester & King, 2009), oder Versorgung großer Lerngruppen bei einmaligem Entwicklungsaufwand (Wulf et al., 2014). Obwohl VLE immer erfolgreicher werden, ist es wichtig Voraussetzungen einer ihrer zentralen Bestandteile,

dem selbstregulierten Lernen, (Testa & Luciano, 2010) zu überprüfen. Vor allem dann, wenn komplexe Aufgaben, wie das Lernen der professionellen Unterrichtswahrnehmung (Sherin & van Es, 2009), dazu gehören. Zentrale Voraussetzungen für das selbstregulierte Lernen sind unter anderem eine hohe Motivation (Chen & Jang, 2010) sowie eine geringe kognitive Belastung (Frederiksen et al., 2020) der Lernenden. Es gibt erste empirische Hinweise darauf, dass diese Faktoren durch Unterstützungselemente wie Video Modeling und Prompts förderbar sind (Renkl, 2014; Schmidt-Weigand et al., 2009). Wir wollten daher diesen Zusammenhang empirisch replizieren und darüber hinaus erfahren, welche Nutzung von Prompts und Video Modeling mit Motivation und kognitiver Belastung einhergehen.

Dafür untersuchten wir 104 Masterstudierende des Grundschullehramts. 51 davon arbeiteten mit einer digitalen Lernumgebung zur Förderung der professionellen Wahrnehmung (abrufbar unter www.prodvis.de) ohne, 53 davon mit digitalen Unterstützungselementen. Die Motivation und die kognitive Belastung wurde während der Bearbeitung jeweils mit entsprechenden Kurzskalen (Wilde et al., 2009; Klepsch et al., 2017) erhoben. Zudem wurden mithilfe eines offenen Antwortformats die Art der Nutzung der digitalen Unterstützungselemente erfasst und anschließend induktiv entsprechende Kategorien gebildet.

Mit Hilfe linearer Regressionsrechnungen konnte herausgefunden werden, dass die Gruppe mit digitalen Unterstützungselementen signifikant motivierter war ($b = .78$, S.E. $= .27$, $p = .004$). Darüber hinaus konnten in der Unterstützungsgruppe - im Gegensatz zum Video Modeling, das keine vergleichbaren Effekte zeigte - inhaltliche Prompts vor allem zur Motivation ($b = .44$, S.E. $= .17$, $p = .012$) und prozedurale Prompts ($b = -1.11$ S.E. $= .34$, $p = .002$) zur Verminderung der kognitiven Belastung beitragen.

Die Bedeutung der Ergebnisse für die Entwicklung weiterer Selbstlernumgebungen unter Verwendung von Unterrichtsvideos und zum Training der professionellen Unterrichtswahrnehmung wird vor dem Hintergrund der Studienlimitationen diskutiert.

D-19: Fachlich-interdisziplinäre Perspektiven auf das Textverstehen: Studien zu Determinanten und Verstehensprozessen im Umgang mit Fachtexten und Aufgaben

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* T03 R03 D75

Fachlich-interdisziplinäre Perspektiven auf das Textverstehen: Studien zu Determinanten und Verstehensprozessen im Umgang mit Fachtexten und Aufgaben

Chair(s): **Anke Schmitz** (Leuphana Universität, Deutschland), **Miriam Dittmar** (PH FHNW)
*Diskutant*in(nen):* **Markus Linnemann** (Universität Koblenz-Landau)

Im Unterricht erfolgt die Wissensvermittlung ob in Printform oder digital weitgehend textbasiert. Entsprechend wird die lernzielbezogene Nutzung von Fachtexten nach wie vor als Schlüsselqualifikation bezeichnet (Weis et al., 2019), die sich auf den Lernerfolg in allen sprachbasierten Unterrichtsfächern auswirkt. Dazu zählen das Verstehen von expositorischen Texten in naturwissenschaftlichen Unterrichtsfächern (Härtig et al., 2015), das Verstehen von Aufgabentexten in Mathematik (Prediger & Krägeloh, 2015) und das reflektierte, kritische Lesen von expositorischen Texten oder literarischen Texten in Deutsch sowie Verfassertexten in Geschichte (Shanahan & Shanahan, 2020).

Zum heutigen Zeitpunkt liegen zahlreiche Arbeiten zum Textverständnis als Produkt und zu Determinanten auf Text- und Leser*innenseite sowie Fördermaßnahmen vor. Grundlegende kognitive Repräsentationen auf der Wort, Satz- und Textebene sind ausführlich beschrieben (Kintsch, 1998; Schnotz, 2006), leser*innenseitige sowie textseitige Prädiktoren des Textverständnisse ermittelt (Cromley & Azevedo, 2007; Schaffner et al., 2004) und eine Reihe von förderlichen Strategieinterventionen entwickelt (Schuster et al., 2020). Fragen nach der Fachlichkeit des Verstehens (Dumas, 2020) wurden bisher allerdings nicht systematisch untersucht. Bis heute ist weitgehend unbekannt, was fachspezifische Verstehensprozesse im Sinne von Disciplinary Literacy (Shanahan & Shanahan, 2020) kennzeichnet. Da ein wechselseitiger Austausch zwischen der psychologischen und fachdidaktischen Forschung zum Textverstehen trotz ihres gemeinsamen Forschungsinteresses bislang wenig stattfand, werden im Symposium interdisziplinäre Studien zum fachlichen Textverstehen vorgestellt, um Spezifika und Generika von Verstehensprozessen, Determinanten und Fördermöglichkeiten auszuloten.

Vorgestellt werden vier Studien, die mit quantitativen und qualitativen Methoden die Fachlichkeit des Textverstehens beleuchten und ihre Ergebnisse mit generischen Befunden und einschlägigen Theorien (bspw. Kintsch, 1998) abgleichen. Es handelt sich um interdisziplinäre Studien unter Beteiligung der Deutschdidaktik, Chemiedidaktik, Geschichtsdidaktik, Mathematikdidaktik, Physikdidaktik und pädagogischer Psychologie, die sich auf das Verstehensprozesse und leser*innenseitige sowie (aufgaben-)textseitige Einflussfaktoren am Beispiel von (Aufgaben-)Texten der Fächer Deutsch, Mathematik, Physik/Chemie und Geschichte konzentrieren. Beitrag 1 fokussiert auf leser*innenseitige Determinanten für das Verstehen von expositorischen Physik- und Chemietexten im Vergleich mit einem narrativen Text, um potenzielle fachspezifische kognitive und motivationale Einflussfaktoren zu isolieren. Beitrag 2 betrachtet expositorische und narrative Verstehensprozesse innerhalb des Faches Deutsch und untersucht mit einer prozessorientierten Analyse die textsortenspezifische Nutzung von kognitiven und metakognitiven Lesestrategien. Beitrag 3 zeigt, wie Verstehensprozesse im Fach Geschichte textseitig und aufgabenseitig so angeleitet

werden können, damit historisches Lernen mit Geschichtslehrmitteltexten besser gelingt. Beitrag 4 untersucht, ob die Platzierung der Fragestellung vor dem mathematischen Aufgabentext als Leseziel das Textverstehen, die Bearbeitungsdauer und die Mathematisierung beeinflusst. In der Diskussion werden die Befunde aus psychologisch-fachdidaktischer Perspektive zusammengefasst, theoretisch verortet und Ansatzpunkte für interdisziplinäre Forschung aufgezeigt.

Beiträge des Symposiums

Kognitive und Motivationale Prädiktoren des Textverstehens in expositorischen versus narrativen Texten

Nadine Cruz-Neri[1], Sascha Bernholt[2], Hendrik Härtig[3], Anke Schmitz[4], Jan Retelsdorf[1]
[1]Universität Hamburg, [2]Leibniz IPN Kiel, [3]Universität Duisburg-Essen, [4]Leuphana Universität Lüneburg

Theoretischer Hintergrund: Expositorische und narrative Texte unterscheiden sich stark hinsichtlich der linguistischen Oberfläche, der Struktur sowie der Voraussetzung von Vorwissen (z. B. Gardner, 2004; Lorch, 2015). Daher ist anzunehmen, dass Schüler:innenmerkmale unterschiedlich bedeutsam für das Textverstehen in expositorischen versus narrativen Texten sind. Bisherige Forschung hat vor allem den Einfluss verschiedener kognitiver Prädiktoren auf das Textverstehen in expositorischen und narrativen Texten untersucht. Es hat sich gezeigt, dass einige kognitive Variablen (z.B. Wortschatz) das Textverstehen in beiden Textgenres vorhersagen, während andere kognitive Variablen (z.B. Lesestrategien) domänenspezifische Prädiktoren sind (z. B. Liebfreund, 2021; Schaffner et al., 2004). Der Effekt der Lesemotivation auf das Textverstehen in Abhängigkeit des Textgenres bleibt jedoch unklar.

Forschungsfragen: Ziel dieser Studie war es zu untersuchen, welche kognitiven und motivationalen Schüler:innenmerkmale domänenübergreifende versus domänenspezifische Prädiktoren für das Textverstehen in expositorischen und narrativen Texten darstellen.

Methode: Die Stichprobe bestand aus N = 261 Achtklässler:innen (34.5% Mädchen; Alter: M = 14.97, SD = 0.82). Um die Fragestellung zu untersuchen, wurden Pfadanalysen gerechnet, in der das Textverstehen in drei expositorischen Texten (in den Fächern Physik und Chemie) sowie in einem narrativen Text als abhängige Variablen fungierten. Unter der Kontrolle von soziodemographischen Variablen (z. B. Migrationshintergrund), wurden die Kompetenz Inferenzen zu bilden (gemessen mit dem Kognitiven Fähigkeitstest; Heller & Perleth, 2000; α = .89), der Wortschatz (gemessen mit dem Kognitiven Fähigkeitstest; Heller & Perleth, 2000; α = .77) und das Lesestrategiewissen (gemessen mit dem Würzburger Lesestrategie-Wissenstest für die Klasse 7–12; Schlagmüller & Schneider, 2007; α = .84) als kognitive Prädiktoren in das Modell aufgenommen. Die Leselust und das Lesen aus Interesse (gemessen mit dem Fragebogen zur habituellen Lesemotivation; Möller & Bonerad, 2007; α = .89 und α = .79) fungierten als motivationale Prädiktoren.

Ergebnisse: Das Textverstehen in expositorischen und narrativen Texten wurde jeweils durch den Wortschatz (β = 0.173–0.381, p < .008) und die Fähigkeit Inferenzen (β = 0.216–0.340, p ≤ .001) zu bilden vorhergesagt. Dementsprechend scheinen diese kognitiven Fähigkeiten domänenübergreifende Prädiktoren für das Textverstehen zu sein. Darüber hinaus sagte das Lesestrategiewissen das Textverstehen jeweils eines expositorischen (β = 0.202, p = .001) und eines narrativen Textes (β = 0.226, p < .001) voraus. Das Lesen aus Interesse war prädiktiv für das

Textverstehen in zwei von drei expositorischen Texten (β1 = 0.156, p = .008; β2 = 0.117, p = .041). Die Leselust war für keinen der Texte prädiktiv für das Textverstehen.

Diskussion: Die Ergebnisse zeigen, dass es sowohl domänenübergreifende als auch domänenspezifische Prädiktoren für das Textverstehen in unterschiedlichen Textgenres und Fachinhalten gibt. Daraus ergeben sich Implikationen wie Schüler:innen angemessen im Schulkontext in ihrem Textverstehen unterstützt werden können. So sollten Lehrkräfte beispielsweise dafür sensibilisiert werden, dass der Wortschatz und die Fähigkeit Inferenzen zu bilden domänenübergreifende Prädiktoren für das Textverstehen sind – und somit einen guten Ansatzpunkt bieten, um das Textverstehen von Schüler:innen zu fördern (z. B. Hall, 2016; Wright & Cervetti, 2016).

Kognitive und metakognitive Lesestrategien. Eine multi-methodale Studie zur Analyse von kognitionspsychologischen Strategien für literarisches Verstehen

Anke Schmitz[1], Wiebke Dannecker[2]
[1]Leuphana Universität Lüneburg, [2]Universität zu Köln

Theoretischer Hintergrund

Die Wirkung von Lesestrategien auf das Verstehen von expositorischen Texten wurde in einer Reihe von Interventionsstudien untersucht. Ergebnis ist, dass sich kognitive und metakognitive Strategien (insbesondere im naturwissenschaftlichen Bereich) positiv auf das fachliche Verstehen und Strategiewissen von Lernenden auswirken (u. a. Schütte et al., 2012; Schuster et al., 2018; Souvignier & Mokhlesgerami, 2006). Allerdings führten Lesestrategietrainings bislang nicht zu Unterschieden im literarischen Verstehen (Mokhlesgerami et al., 2007; Trenk Hinterberger, 2006). Eine empirische Spezifizierung von kognitiven und metakognitiven Lesestrategien für das Verstehen von literarischen Texten steht bislang aus. Da literarisches Verstehen im Vergleich mit dem expositorischen Verstehen die Akzeptanz eines mehrschichtigen Situationsmodells, das Aushalten von Mehrdeutigkeit und eine Aufmerksamkeit für die sprachliche und ästhetische Gestaltung des Textes erfordert (Kintsch, 1994; Zwaan, 1996), ist anzunehmen, dass Lesestrategien, die auf dem Konzept informed strategies for learning beruhen, nur bedingt zu literarischem Lesen passen. Vermutlich fordert das Verstehen literarischer Texte literarisch-orientierte Strategien und Fähigkeiten (Kintsch, 1994), Textelemente, die sich nicht direkt in ein Situationsmodell integrieren lassen, im Leseprozess nicht zu vernachlässigen, sondern sie als internen Widerspruch im Gedächtnis zu behalten, eher emotionale Aspekte als inhaltliches Vorwissen in das Situationsmodell zu integrieren, und die sprachliche Gestaltung und Darstellungsintention in Konstruktion des Situationsmodells einzubeziehen.

Fragestellungen

Im Beitrag wird eine multi-methodale Studie vorgestellt, in das strategische Lesen eines literarischen Text mit dem Lesen eines expositorischen Textes kontrastiert wird. Untersucht wird (1), welche kognitiven und metakognitiven Lesestrategien Leser*innen beim Lesen eines literarischen Textes im Vergleich mit einem expositorischen Text im Leseprozess verwenden und berichten und (2) in welchem Zusammenhang der Strategieeinsatz mit dem literarischen Textverständnis steht.

Methode

An der Studie nahmen 54 Lehramtsstudierende teil (23 Jahre; 95% weiblich, 5% männlich), die einen literarischen Text und einen expositorischen Text (within-subjects) lasen. Lesestrategien wurden durch text traces (Markierungen, Anmerkungen) während des Lesens und retrospektive Interviewfragen erhoben, das jeweilige Textverständnis über Notizen zu beiden Texten erfasst. Die text traces wurden hinsichtlich des Umfangs und ihrer Zielgerichtetheit angesichts zentraler, für das Verständnis relevanter Aspekte (max. 11 pro Text), kodiert (Scott, 2008). Die Interviewantworten wurden hinsichtlich kognitiver Lesestrategien (Organisation, Elaboration, Wiederholung) und metakognitiver Strategien induktiv-deduktiv kodiert. Das Textverständnis wurde anhand der Kohärenz der Notizen auf einer Skala von 1-5 kodiert (1 = oberflächenbezogene inkohärente Notizen, 5 = inhaltlich kohärente, über den Text reflektierende Notizen) (Bråten & Samuelstuen, 2007). Alle Kodierungen wurden von drei geschulten Wissenschaftler*innen vorgenommen (Interrater-Reliabilität κ = .75-95).

Ergebnisse und Ausblick

Die gepaarten t-Tests zeigen, dass sich die strategischen Zugänge within-subjects unterscheiden. Beim literarischen Text wurde weniger, aber gezielter markiert, als im expositorischen Text, wo es mehr Markierungen gab, aber unfokussiert vorgegangen wurde. Das über die Notizen erfasste Textverständnis zum literarischen Text spiegelte sich in signifikant umfangreicheren sowie kohärenteren Notizen wider. Die Selbstberichte verdeutlichen, dass die Proband*innen mehr und (spezifisch literarische) Elaborationen beim Lesen des literarischen Texts berichteten ($t(1, 53) = 3.523$, $p < .001$). Die Korrelationsanalysen zwischen dem jeweiligen Textverständnis, den text traces sowie den selbstberichteten kognitiven und metakognitiven Lesestrategien zeigen unterschiedliche Muster: Beim literarischen Lesen gehen metakognitive Lesestrategien mit besserem Textverständnis ($r = .286$*), mehr Elaborations- ($r = .315$*) und Organisationsstrategien ($r = .344$*) einher; beim expositorischen Text hingegen liegen nur Korrelationen zwischen Metakognition und Wiederholungs- ($r = .310$*) und Elaborationsstrategien ($r = .440$*) vor.

Diskutieren lassen sich die Spezifika der Lesestrategien beim Lesen des literarischen Textes und in welcher Relation sie zu Strategietrainings stehen (Mokhlesgerami et al., 2007). Ferner lässt sich das Design einer geplanten Interventionsstudie diskutieren, um Lernenden, die Schwierigkeiten mit dem Verstehen literarischer Texte haben (Stark, 2017), Lesestrategien für das Verstehen von literarischen Texten zu vermitteln.

Unterstützung des Textverstehens im Fach Geschichte

Miriam Dittmar, Sandro Brändli, Claudia Schmellentin
PH FHNW

Theoretischer Hintergrund

Historisches Wissen wird in hohem Maß mithilfe von Texten vermittelt und angeeignet (Rüsen, 2008). Die Anforderungen, die die Geschichtslehrmitteltexte an die literalen Kompetenzen der Jugendlichen auf der Sekundarstufe I stellen, übersteigen jedoch häufig die vorhandenen Kompetenzen (von Borries et al., 2005). So erfordern Geschichtslehrmitteltexte nicht einfach die Informationsentnahme, sondern das kritische Hinterfragen von Narrativen (Wineburg, 2001) oder das Anwenden komplexer historischer Denkoperationen, wie das Erkennen von Chronologien, Ursache-Folge-Beziehungen

oder die Belegbarkeit von historischen Aussagen durch Quellen einzuschätzen (Kühberger, 2012). Das bedeutet, die Textverstehenskompetenzen sind mindestens der PISA-Kompetenzstufe V zuzuordnen (Weis et al. 2019). Um diese hierarchiehohen Textverstehensleistungen zu erreichen, sind fachspezifische Lesestrategien notwendig, an welche die Jugendlichen der Sekundarstufe I allerdings kaum hingeführt werden (Handro, 2018). Auch besteht nur wenig empirisch fundiertes Wissen darüber, wie diese Hinführung für die Sekundarstufe I didaktisch zu modellieren ist: Textseitig schlägt Schrader (2020) vor, den Textverstehensprozess mittels an geschichtsdidaktischen Zielen orientierte Kohärenzbildungshilfen (Schnotz, 1994) zu entlasten, welche die globalen Thema-Strukturen, übergeordneten Wissenseinheiten, funktionalen Relationen sowie die Leseziele explizieren. Aufgabenseitig kann der Leseprozess sowie die Anwendung der fachspezifischen Lesestrategien angeleitet werden. (Lindauer el al., 2013, Perry et al., 2008).

Forschungsziel und Fragestellung

Ziel des im Symposium vorzustellenden SNF-geförderten Forschungsprojektes ist es daher, empirisch fundiertes Wissen darüber zu generieren, wie Textverstehensprozesse auf der Sekundarstufe I im Fach Geschichte lesedidaktisch unterstützt werden können und so das Lernen mit Texten allen Lernenden zugänglich gemacht wird. Daten liegen dazu aus einer qualitativen Teilstudie vor, welche folgende Forschungsfrage bearbeitet: Wie verstehen Achtklässler*innen Geschichtslehrmitteltexte, welche mit textseitigen lesedidaktischen Maßnahmen (Kohärenzbildungshilfen) angereichert sind, welche Verstehensschwierigkeiten treten dabei trotz der lesedidaktischen Unterstützung auf und wie kann diesen Schwierigkeiten mit aufgabenseitigen lesedidaktischen Maßnahmen begegnet werden?

Methodisches Vorgehen

Die Fragestellung wurde mithilfe von qualitativen Leseprozessbeobachtungen untersucht. Dazu wurde zunächst ein gut verständlicher Geschichtslehrmitteltext mit lesedidaktisch orientierten Kohärenzbildungshilfen entwickelt, welcher dann in Leseprozessbeobachtungen eingesetzt wurde. An den Leseprozessbeobachtungen nahmen 10 Schüler*innen aus achten Sekundarklassen (nicht-gymnasial, 14 Jahre, 60% weiblich, 40% männlich) in der Schweiz teil. Nach einer Stillesephase und der Bearbeitung von schriftlichen Fragen zu den Textinhalten wurden die Schüler*innen mittels offener und geschlossener Fragen zu ihrem Leseprozess und dem Verständnis von ausgewählten Textstellen befragt, indem sie ihre Antworten zu den Fragen mit entsprechenden Textstellen begründen mussten (fokussiertes Interview mit Stimulated Recall, Friebertshäuser & Langer, 2010). Die Begründungen der Schüler*innen wurden hinsichtlich Textbezogenheit (textnahe oder textferne Argumentation), Verarbeitungstiefe mit Bezug zu den Pisa-Kompetenzstufen (Pisa I/II: Verarbeitung auf Satzebene, z.B. wörtliche Bedeutung erfassen, Pisa III/IV: lokale Kohärenzbildung, z.B. Verknüpfen von Textstellen, Weis et al. 2019) und der Passung von schülerseitiger Begründung und den Anforderungen des Frageitems (problematisch, unproblematisch) deduktiv kodiert. Als problematische Begründungen galten hier z.B. textnahe Begründung, die die falsche Textstelle fokussierten oder rein vorwissensgeleitete Begründungen. Die problematischen Schüler*innenbegründungen wurden anschließend inhaltsanalytisch kategorisiert. So kann auf

standardisierte Weise erkannt werden, welche Textstellen bzw. Sachverhalte besondere Verstehensprobleme bereiten, die einer weiteren lesedidaktischen Unterstützung bedürfen.

Ergebnisse

Insgesamt liegen aus den 10 Interviews n = 230 Begründungen der Schüler*innen aus 32 Frageitems vor. Die Ergebnisse aus den qualitativen Leseprozessbeobachtungen zeigen, dass trotz der Aufbereitung des Textes durch lesedidaktische Kohärenzbildungshilfen immer noch 44% der Begründungen als problematisch eingestuft werden müssen. Dieser Prozentsatz erhöht sich auf 50%, wenn man nur die Frageitems einberechnet, die eine vertieftere Verarbeitung von Textabschnitten erfordern (ab PISA-Kompetenzstufe III). Schaut man Items an, die komplexe historische Denkoperationen erfordern, so erhöht sich der Prozentsatz der problematischen Begründungen noch einmal auf 60-75%. Die problematischen Begründungen der Schüler*innen lassen sich inhaltsanalytisch in sechs Kategorien zusammenfassen, welche Prinzipien zu aufgabenseitigen lesedidaktischen Maßnahmen generieren. In einer nächsten Teilstudie soll nun die aufgabenseitige Verstehensprozesssteuerung mithilfe einer quantitativen Interventionsstudie im between-subjects Design auf ihre Wirksamkeit hin untersucht werden.

Zur Rolle der Fragestellung in mathematischen Modellierungsaufgaben: Wirkt sich das Voranstellen der Fragestellung vor den Aufgabentext positiv auf das Textverstehen aus?

Valentin Böswald, Stanislaw Schukajlow
Universität Münster

Theorie

Interdisziplinäre Grundlage des Verstehens von Texten ist das Lesen im Sinne der Verarbeitung von visuellen Reizen wie etwa Buchstaben und Wörtern. Eine wichtige Fachkompetenz im Fach Mathematik ist die Fähigkeit, realitätsbezogene Aufgaben (Modellierungsaufgaben) zu bearbeiten. Eine zentrale Aktivität ist hierbei die Konstruktion eines mathematischen Modells (Mathematisieren; Niss & Blum, 2020). Das Voranstellen der Fragestellung vor den Aufgabentext sollte das Leseziel spezifizieren und das Textverstehens beim mathematischen Modellieren unterstützen (Böswald & Schukajlow, 2022; Thevenot et al., 2007). Das Textverstehen hilft bei der Selektion und Gewichtung von mathematisch bzw. situational relevanten Informationen und begünstigt das Mathematisieren (Leiss et al., 2019). Der Leseprozess sollte effizienter verlaufen, wenn die Fragestellung vorangestellt wird, d.h. zeitökonomischer (geringere Bearbeitungsdauer) bei gleichbleibender Leistung. Zudem könnten Mediationseffekte vom Voranstellen der Fragestellung auf das Mathematisieren über das Textverstehen existieren.

Fragestellungen

1. Beeinflusst die Position der Fragestellung a) das Textverstehen von mathematischen Modellierungsaufgaben b) Bearbeitungsdauer und c) die Mathematisierung?

2. Mediiert das Textverstehen den Effekt zwischen der Position der Fragestellung und die Mathematisierung?

3. Inwiefern beeinflusst die Position der Fragestellung das Leseverhalten beim Textverstehen von mathematischen Modellierungsaufgaben?

Methode

192 Schüler:innen (43.8% weiblich, M Alter = 16.0 Jahre) der Jahrgangsstufen 10-12 wurden randomisiert der Experimentalgruppe (EG, Fragestellung vor dem Text) oder der Kontrollgruppe (KG, Fragestellung nach dem Text) zugewiesen. In der EG wurde zuerst die Fragestellung und dann der Aufgabentext eingeblendet. Als Messinstrumente (alle computerbasiert) wurden ein Mathematisierungstest (4 Items) sowie ein Test zur Erfassung des Textverstehens bei Modellierungsaufgaben (verbal und piktorial, 48 Items) eingesetzt und die Bearbeitungsdauer erfasst. In einer zweiten Stichprobe (75 Schüler:innen, 54.7% weiblich, M Alter = 16.9 Jahre) wurden zur Beantwortung der dritten Forschungsfrage zusätzlich die Blickbewegungen beim Textverstehenstest aufgezeichnet. Als Kontrollvariablen wurden die Lesegenauigkeit, die Lesegeschwindigkeit sowie das mathematische Vorwissen erhoben. Alle Skalen zeigten gute Reliabilität und die Pfadmodelle hatten passende Fitwerte.

Ergebnisse und Diskussion

In einer Pfadanalyse wurden zwei Pfadmodelle getestet, je eins mit dem verbalen und piktorialen Textverstehenstest. Es zeigte sich, dass die Position der Fragestellung kein signifikanter Prädiktor für das Textverstehen oder die Mathematisierung – weder direkt, noch total oder indirekt – ist. Die Bearbeitungsdauer wird wie erwartet signifikant negativ von der Position der Fragestellung beeinflusst ($by = -.283$, $SE = .141$, $p < .05$), sodass eine Effizienzsteigerung geschlussfolgert werden kann.

Mithilfe von einfaktoriellen Varianzanalysen wurden die Mittelwerte der Blickbewegungsparameter beim Textverstehen verglichen. Im Hinblick auf die globalen Blickbewegungsmaße durchschnittliche Fixationsdauer, Wörter pro Fixationsminute, Sakkaden pro Wort sowie Regressionen pro Sakkade konnten keine signifikanten Unterschiede festgestellt werden. Hinsichtlich der mathematisch und situational relevanten Informationen in den Aufgabentexten wurden bei den lokalen Blickbewegungsmaßen signifikante Unterschiede in der totalen Verweilzeit festgestellt (in Sekunden; $MEG = 75.25$, $SD = 22.99$; $MKG = 96.31$, $SD = 35.73$; $F = 9.265$, $p < .003$, $\eta 2 = .11$). Kein signifikanter Unterschied zeigte sich bezüglich der Verweildauer auf irrelevanten Informationen ($MEG = 62.38$, $SD = 21.10$; $MKG = 72.27$, $SD = 25.85$; $F = 3.307$, $p = .073$, $\eta 2 = .04$). Die kürzere Verweilzeit auf relevanten Informationen in der Experimentalgruppe kann als Indiz für die Wirksamkeit der Voranstellung der Fragestellung auf die Effizienz und als Bestätigung der hypothesenkonformen Befunde auf die Bearbeitungsdauer in der Pfadanalyse gedeutet werden. Die Unterschiede im Leseverhalten deuten darauf hin, dass die fehlenden Effekte des Voranstellens der Fragestellung auf das Textverstehen und Mathematisieren durch die geänderten Vorgehensweisen der Schüler:innen beim Lesen erklärt werden können.

Die Ergebnisse weisen auf theoretischer Ebene darauf hin, dass das Spezifizieren des Leseziels einen Einfluss auf Leseprozesse hat (Böswald & Schukajlow, 2022; Thevenot et al., 2007). Eine praktische Implikation ist die Möglichkeit, den Leseprozess und die Bearbeitungszeit durch die Position der Fragestellung bei der Aufgabengestaltung zu beeinflussen.

D-20: Die Bedeutung von aktiver Schüler*innenbeteiligung im Unterricht: Individuelle Lernvoraussetzungen, kontextuelle Determinanten und der Zusammenhang mit Lernerfolg

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* R11 T00 D05

Die Bedeutung von aktiver Schüler*innenbeteiligung im Unterricht: Individuelle Lernvoraussetzungen, kontextuelle Determinanten und der Zusammenhang mit Lernerfolg

Chair(s): **Ricardo Böheim** (Technische Universität München), **Tina Seidel** (Technische Universität München)
*Diskutant*in(nen):* **Doris Holzberger** (Technische Universität München)

Vor dem Hintergrund empirischer Befunde und sozial-konstruktivistischer Lerntheorien besteht mittlerweile ein breiter Konsens darüber, dass die aktive Schüler*innenbeteiligung im Unterricht eine zentrale Rolle für erfolgreiche Lehr-Lernprozesse spielt (Chi & Wylie, 2014). Produktive Lehrkraft-Schüler*innen-Interaktionen und ko-konstruktives Lernen setzen voraus, dass sich Schüler*innen aktiv im Unterrichtsgeschehen einbringen (Resnick et al., 2015). In der Unterrichtsforschung wird die aktive Schüler*innenbeteiligung häufig über beobachtbare Indikatoren wie Meldungen oder Wortbeiträge konzeptualisiert, die üblicherweise auf der Basis von Unterrichtsvideos erfasst werden. Obwohl zahlreiche Publikationen zum Konstrukt des „student engagement“ vorliegen (Sinatra et al., 2015), werden beobachtbare Verhaltensindikatoren der Schüler*innenbeteiligung vergleichsweise selten untersucht (Böheim et al., 2020). Vor diesem Hintergrund bleiben zentrale Fragen nach Bedingungen und Wirkungen beobachtbarer Schüler*innenbeteiligung weitestgehend unbeantwortet.

Ziel des Symposiums ist es, empirische Befunde zur Schüler*innenbeteiligung vorzustellen und dabei wichtige Fragen zur Stabilität und Verteilung, zu individuellen und kontextabhängigen Determinanten sowie zur Bedeutung für den Lernerfolg zu klären. In allen Beiträgen wird aktive Schüler*innenbeteiligung mittels Videoanalyse und anhand beobachteter Verhaltensindikatoren erfasst. Die Fragestellungen der Beiträge werden mittels quantitativer Analysemethoden untersucht. Die Stichproben umfassen sowohl die Primar- als auch die Sekundarstufe und wurden in unterschiedlichen Schulfächern erhoben.

Im ersten Beitrag werden die Häufigkeit und Verteilung der Beteiligung von Schüler*innen im Deutschunterricht untersucht. Dabei wird zwischen einer schüler*innengesteuerten (Meldungen) und einer lehrkraftgesteuerten (Aufrufe durch die Lehrkraft) Beteiligung differenziert. Es zeigt sich, dass die Variation im Meldeverhalten im Zusammenhang mit individuellen Lernvoraussetzungen steht. Die Ergebnisse deuten zudem auf die Problematik hin, dass Lehrpersonen im öffentlichen Unterrichtsgespräch vor allem mit leistungsstarken Schüler*innen interagieren.

Auch im zweiten Beitrag wird zwischen einer schüler*innengesteuerten und einer lehrkraftgesteuerten Beteiligung im Unterricht unterschieden. Die in der Primar- und Sekundarstufe durchgeführten Analysen zeigen, dass vor allem der schüler*innengesteuerten Beteiligung, insbesondere den Meldungen, eine entscheidende Rolle für die Leistungsentwicklung zugeschrieben werden kann.

Im dritten Beitrag werden Schüler*innen-Meldungen im Verlauf des Schuljahres mittels eines Längsschnittdesigns untersucht. Es zeigt sich, dass die Meldehäufigkeit in der Sekundarstufe ein relativ stabiles Schüler*innenmerkmal ist. Darüber hinaus wird gezeigt, dass das Meldeverhalten eine Voraussetzung aber auch ein Outcome von motivationalen und kognitiven Lernvoraussetzungen sein kann.

Im vierten Beitrag wird eine prozessbezogene, situative Analyse zur Schüler*innenbeteiligung durchgeführt. Die Ergebnisse unterstreichen die zentrale Rolle von Lehrpersonen in unterrichtlichen Interaktionen und zeigen einen deutlichen Zusammenhang zwischen der Art bzw. Qualität der Unterrichtsgestaltung und der Schüler*innenbeteiligung.

Zum Abschluss werden die Beiträge von einer ausgewiesenen Expertin im Bereich der Forschung zu Interaktionen im Unterricht diskutiert, die Bedeutung der Ergebnisse hervorgehoben und weitere Forschungsperspektiven aufgezeigt.

Beiträge des Symposiums

Wer hat im Anfangsunterricht der Grundschule was zu melden? Teilhabe am Unterrichtsgespräch im Deutschunterricht des ersten Schuljahres

Ann-Katrin Denn[1], Miriam Hess[2], Frank Lipowsky[3]
[1]Hochschule für öffentliches Management und Sicherheit, [2]Universität Bamberg, [3]Universität Kassel

Theoretischer Hintergrund und Forschungsstand

Die Beteiligung von Kindern am Unterrichtsgespräch wird als eine bedeutsame Facette des Interaktionsgeschehens im Unterricht gesehen (Mercer&Dawes, 2014; Pauli&Reusser, 2015). Dabei kann die Beteiligung am Unterricht in schüler- und lehrkraftgesteuerte Facetten unterteilt werden: Meldungen gelten als schülergesteuerte und Aufrufe der Lehrperson als lehrkraftgesteuerte Facette der Unterrichtsbeteiligung (Sacher, 1995). Welche Kinder sich im Unterricht melden und welche aufgerufen werden, wurde bereits in mehreren Studien im Grundschulunterricht untersucht, in denen sich zeigte, dass Kinder mit vergleichsweise günstigen Voraussetzungen häufiger am Unterrichtsgespräch partizipierten (Jones, Dindia&Tye, 2006; Decristan et al., 2020; Denn, 2021). Inwiefern diese Ergebnisse auch für den Anfangsunterricht im Fach Deutsch bestätigt werden können, soll in diesem Beitrag untersucht werden.

Fragestellungen

1) Wie häufig melden sich Kinder und wie häufig werden sie von der Lehrperson aufgerufen, wenn sie sich gemeldet haben?

2) Wie häufig werden Kinder ohne Meldung im Unterricht aufgerufen?

3) Welche Kinder melden sich und welche werden aufgerufen (in Abhängigkeit von Geschlecht, Intelligenz, Konzentration und Vorläuferfähigkeiten im Fach Deutsch)?

Methode

Projekt und Stichprobe

Die Daten stammen aus dem Projekt PERLE, in dem die Persönlichkeits- und Lernentwicklung von Grundschulkindern über einen Zeitraum von vier Jahren untersucht wurde und das Videostudien in drei Fächern beinhaltete (Lipowsky, Faust&Kastens, 2013). Die Videografie des Deutschunterrichts fand in der Mitte des ersten Schuljahres statt. Die Stichprobe dieses Beitrags besteht aus 48 Unterrichtsvideos mit 640 Kindern (Min=5, Max=25, M=13.33, SD=6.09) und einem Unterrichtsgespräch von durchschnittlich 43.19 Minuten (Min=10.83, Max=77.00, SD=16.99).

Kodierung der Meldungen und Aufrufe

Das Kodiersystem der Meldungen und Aufrufe besteht aus vier Kategorien „Meldung mit Aufruf“, „Meldung ohne Aufruf“, „Aufruf ohne Meldung“ und „Aufruf – Meldung nicht erkennbar“ (Hess&Denn, 2019). Die Übereinstimmungswerte liegen über den gesetzten Mindestkriterien (ebd.).

Unabhängige Variablen

Das Geschlecht der Kinder wurde als unabhängige dichotome Variable in die Analysen einbezogen (Jungen=0, Mädchen=1).

Die nachfolgenden Variablen wurden zu Beginn des ersten Schuljahres erhoben: Die schriftsprachlichen Vorläuferfertigkeiten (Rohwerte Silben klatschen: Min=0.00, Max=8.00, M=6.64, SD=1.62; Rohwerte Anlaute hören: Min=0.00, Max=8.00, M=5.73, SD=2.20), die Konzentration mit der Anzahl der richtigen Lösungen als Variable (Min=0.00, Max=23.00, M=14.02, SD=4.47) und die Grundintelligenz mit dem CFT 20-R (Cattell, Weiß&Osterland, 1997) (M=114.51; Min=80.50; Max=146.50; SD=11.35).

Das Selbstkonzept im Lesen wurde in der Mitte des ersten Schulhalbjahres mit einem selbst entwickelten Fragebogen erfasst (α=.80, M=2.35, SD=0.55; Poloczek, Greb&Lipowsky, 2009).

Analysen

Die Analysen werden als lineare Regressionen mit dem Programm SPSS durchgeführt.

Ergebnisse und ihre Bedeutung

Im Mittel melden sich die Kinder 0.19 Mal pro Minute im klassenöffentlichen Unterricht (Min=0.00, Max=1.27, SD=0.21), wobei zu berücksichtigen ist, dass sich 12% der Grundschulkinder während des Unterrichts überhaupt nicht melden.

In den Regressionsanalysen zeigt sich, dass sich Mädchen tendenziell etwas häufiger melden als Jungen. Des Weiteren melden sich Kinder umso häufiger, je besser ihr Selbstkonzept im Lesen und die Vorläuferfertigkeit „Anlaute hören“ ausgeprägt sind (Selbstkonzept Lesen:ß=.13; p≤.01; Anlaute hören:ß=.10; p≤.05).

Betrachtet man die Aufrufe mit vorangegangener Meldung, werden im Mittel 28.13% der Meldungen aufgerufen (Min=0%, Max=100%, SD=25.72%). Regressionsanalysen zeigen, dass die Lehrpersonen aus dem Angebot der sich meldenden Kinder eher die leistungs- und selbstkonzeptstärkeren Kinder auswählen und aufrufen (Selbstkonzept Lesen:ß=.10; p≤.05; Anlaute hören:ß=.08; p=.071).

Betrachtet man die lehrkraftgesteuerte Beteiligung, also die Aufrufe, denen keine Meldung eines Kindes vorangegangen ist, zeigt sich, dass diese von allen Aufrufen eines Kindes im Mittel 25.29% ausmachen (Min=0%, Max=100%, SD=32.19%). Durch die Regressionsanalysen wird deutlich, dass Lehrpersonen bei Aufrufen ohne vorangegangene Meldung vor allem Jungen aufrufen und Kinder, die niedrige Werte in der Konzentration aufweisen (Geschlecht:ß=-.09; p≤.05; Konzentration:ß=-.13; p≤.01).

Es können die Ergebnisse anderer Studien bestätigt werden, die zeigten, dass Lehrpersonen vermehrt leistungsstärkere Kinder im Grundschulunterricht partizipieren lassen (Decristan et al., 2020; Denn, 2021). Diskutiert werden soll u.a. die Bedeutung von lehrkraftinitiierten Aufrufen bei Kindern.

Schüler*innengesteuerte und Lehrkraftgesteuerte Beteiligung von Schüler*innen am Unterrichtsgespräch in Primar- und Sekundarstufe

Nina Jansen[1], Jasmin Decristan[1], Benjamin Fauth[2]
[1]Bergische Universität Wuppertal, [2]Universität Tübingen, Institut für Bildungsanalysen Baden-Württemberg (IBBW)

Die Beteiligung der Schüler*innen am Unterrichtgespräch gilt als zentraler Bestandteil des Lernens und ist auch gegenwärtig Gegenstand der empirischen Unterrichtsforschung (z. B. Böheim et al., 2020; Schnitzler et al., 2021). Empirisch lassen sich Effekte der Schüler*innenbeteiligung am Unterrichtsgespräch auf fachliche Lernleistungen sowohl in der Grundschule (z. B. Ing et al., 2015) als auch in der Sekundarstufe (Schnitzler et al., 2021) aufzeigen, jedoch ist die Forschungslage hierzu nicht konsistent. Theoretisch lassen sich zwei Perspektiven zur Erklärung des Zusammenhangs zwischen Schüler*innenbeteiligung und fachlichen Leistungen ausmachen: Erstens kann die Beteiligung als ein Indikator für das Engagement der Schüler*innen beim Lernen betrachtet werden (z. B. Fredricks et al., 2016). Eine höhere Unterrichtsbeteiligung wird dabei vor allem mit mehr behavioralem Engagement in Verbindung gebracht (z.B. Troll et al., 2020). Zweitens wird die Rolle der verbalen Interaktion von Schüler*innen mit der Lehrkraft adressiert und hierfür unter Rückbezug auf soziokulturelle Theorien auf die Bedeutung des gemeinsamen Diskurses zum Fachinhalt verwiesen (z.B. Sedova et al., 2019). Die meisten empirischen Studien zur Beteiligung von Schüler*innen am Unterrichtsgespräch beziehen sich jedoch entweder auf die erste (Engagement; wir operationalisieren dies über „schüler*innengesteuerte Beteiligung“) oder die zweite (gemeinsame verbale Interaktion mit der Lehrkraft; „lehrkraftgesteuerte Beteiligung“) Perspektive. Vor dem Hintergrund der damit einhergehenden unterschiedlichen Implikationen für Forschung und Praxis sollen in diesem Beitrag die beiden Forschungsstränge systematisch aufgegriffen und deren jeweiliger Erklärungsbeitrag für das fachliche Lernen empirisch geprüft werden. Um die Robustheit der Ergebnisse zu prüfen, soll sowohl die Primarstufe als auch die Sekundarstufe einbezogen werden.

Vor diesem Hintergrund werden im Beitrag die folgenden beiden Forschungsfragen verfolgt:

1. Welche Bedeutung hat die schüler*innengesteuerte Unterrichtsbeteiligung für Leistungen in der Primar- und Sekundarstufe?

2. Welche Bedeutung hat die lehrkraftgesteuerte Unterrichtsbeteiligung für Leistungen in der Primar- und Sekundarstufe?

Es wurden Daten aus zwei Videostudien mit standardisierten Leistungstests zu zwei Erhebungszeitpunkten herangezogen. Der erste Datensatz bezieht sich auf die Primarstufe (3. Jahrgangsstufe, naturwissenschaftlicher Sachunterricht; Hardy et al., 2011) mit 35 Klassen und 681 Schüler*innen, der zweite Datensatz auf die Sekundarstufe I (8. bis 10. Jahrgangsstufe, Mathematik; TALIS Videostudie Deutschland, Grünkorn et al., 2020) mit 40 Klassen und 932 Schüler*innen. Zur Erfassung der Unterrichtsbeteiligung wurde in jeder Klasse ein ca. 10-minütiger Videoausschnitt des klassenöffentlichen Unterrichtsgesprächs ausgewählt und die Unterrichtsbeteiligung jeder Schülerin und jedes Schülers anhand von Meldungen sowie mündlichen Beiträgen mit und ohne Meldung kodiert (vgl. Schema von Sacher, 1995). Die Indikatoren wurden anschließend zu schüler*innengesteuerter und lehrkraftgesteuerter Unterrichtsbeteiligung zusammengefasst. Insgesamt wurden so in der Primarstufe im Mittel 3.54 schüler*inngengesteuerte und 1.13

lehrkraftgesteuerte Beteiligungen kodiert, in der Sekundarstufe waren es im Mittel 5.22 und 1.04 Beteiligungen. Um die Fragestellungen zu prüfen, wurden Mehrebenenanalysen mit der Posttest-Leistung als abhängige Variable herangezogen. Als Prädiktoren wurde die Häufigkeit der schüler*inngengesteuerten Beteiligung oder der lehrkraftgesteuerten Beteiligung verwendet und in einem zweiten Schritt jeweils zusätzlich für zentrale Lernvoraussetzungen (inkl. Prätest) kontrolliert.

Die Ergebnisse dieses Beitrags zeigen signifikante Zusammenhänge sowohl von der Häufigkeit der schüler*inngengesteuerten Beteiligung als auch von der Häufigkeit der lehrkraftgesteuerten Beteiligung mit der Posttestleistung in Primar- und Sekundarstufe. Nach Kontrolle von Lernvoraussetzungen resultierte in der Sekundarstufe nur noch für die Häufigkeit von schüler*inngengesteuerter Beteiligung ein signifikanter Effekt auf die Mathematikleistung. Ergänzende Analysen zeigen den eigenständigen Erklärungsbeitrag von Schüler*innen-Meldungen für Leistungen in den beiden Studien aus Primar- und Sekundarstufe auf. Insgesamt zeigen die Ergebnisse somit vor allem die Bedeutung des Engagements von Schüler*innen im Unterricht auf. Das Potenzial der verbalen Interaktion zwischen Lehrkraft und Schüler*innen für die weitere Anregung von Lernprozessen scheint hingegen bislang nicht hinreichend genutzt zu werden.

Die Stabilität von Schülermeldungen und deren reziproker Zusammenhang mit Selbstkonzept und kognitiver Lernaktivität im Verlauf des Schuljahres

Ricardo Böheim[1], Martin Daumiller[2], Tina Seidel[1]
[1]Technische Universität München, [2]Universität Augsburg

Theoretischer Hintergrund

Schülermeldungen spielen eine zentrale Rolle für die aktive Beteiligung am Unterrichtsgespräch (Decristan et al., 2019). Bevor Schüler*innen einen Unterrichtsbeitrag machen können, müssen sie sich in aller Regel melden (Dixon et al., 2009). Aktuelle Forschungsstudien unterstreichen die Bedeutung von Schülermeldungen und zeigen positive Zusammenhänge mit motivationalen und kognitiven Lernermerkmalen (Jansen et al., 2022; Schnitzler et al., 2021). Allerdings stützen sich bisherige Erkenntnisse größtenteils auf Querschnittsstudien mit Unterrichtsbeobachtungen aus nur einer Schulstunde. Wichtige Fragen zur zeitlichen Stabilität dieses Verhaltens sowie zu den reziproken Zusammenhängen mit Lernermerkmalen im Verlauf des Schuljahres sind daher noch ungeklärt.

Das erste Ziel des Beitrages ist es, die Stabilität von Schülermeldungen über mehrere Schulstunden zu untersuchen. Befunde zur Unterrichtsbeteiligung von Schüler*innen zeigen eine hohe zeitliche Stabilität, wobei Klassen- und Lehrkraftmerkmale üblicherweise nur einen geringen Anteil der Variabilität zwischen Schüler*innen erklären (Skinner & Pitzer, 2012). Da Schülermeldungen ein eher situationsspezifisches Verhalten darstellen, wird hingegen eine geringere Stabilität über Schulstunden erwartet. Außerdem wird davon ausgegangen, dass Unterschiede zwischen Lehrkräften eine deutlich größere Bedeutung bei der Erklärung der Variabilität spielen, da Meldungen zu einer unmittelbaren Interaktion mit Lehrkräften führen.

Das zweite Ziel des Beitrags ist es, die reziproken Beziehungen zwischen Schülermeldungen, dem Selbstkonzept und der kognitiven Lernaktivität von Schüler*innen zu untersuchen. Basierend auf bisheriger Theorie und Forschung zu Angebots-Nutzungs-Modellen von Unterricht werden

motivationale und kognitive Lernermerkmale als Prädiktoren von Unterrichtsverhalten verstanden (Brühwiler & Blatchford, 2011). Allerdings könnten auch Schülermeldungen selbst Veränderungen in Lernermerkmalen hervorrufen. Schüler*innen, die sich melden, erhalten in der Regel Rückmeldung zu ihrem Beitrag. Diese könnten sich einerseits positiv auf die Entwicklung des Selbstkonzepts auswirken (z.B. durch positives Lehrerfeedback; Dresel & Ziegler, 2006) und gelichzeitig die kognitive Lernaktivität fördern (z.B. durch Ko-konstruktion von Denkprozessen; Chi & Wylie, 2014).

Methode

An der Studie nahmen 376 Schüler*innen aus 19 unterschiedlichen Klassen der Sekundarstufe teil. Zur Erfassung der Meldehäufigkeit wurden alle Klassen zu Beginn und zum Ende des Schuljahres während einer Schulstunde videographiert. Bei 10 Klassen wurden zusätzlich zwei weitere Videoaufnahmen in der Mitte des Schuljahres durchgeführt. Die Kodierung der Schülermeldungen erfolgte mittels der Software Interact (Interrater-Reliabilität $\kappa = .69$). Das akademische Selbstkonzept, sowie die kognitive Lernaktivität der Schüler*innen wurde mittels Fragebogen mit jeweils 5 Items zu Beginn und zum Ende des Schuljahres erhoben (Marsh et al., 2005; OECD, 2005).

Zur Untersuchung der Stabilität von Schülermeldungen wurden Videodaten aus allen vier Messzeitpunkten (insgesamt 58 Schulstunden) herangezogen. Zur Bestimmung der Varianzaufteilung zwischen Schüler*innen, Lehrkräften und Schulstunden wurden Intraklassenkorrelationen (ICCs) mittels eines Cross-Classified Random Effects Models berechnet. Die reziproken Beziehungen zwischen Schülermeldungen, Selbstkonzept und kognitiver Lernaktivität wurden in einer Cross-Lagged-Panel-Analyse unter Berücksichtigung der genesteten Datenstruktur untersucht.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass ca. die Hälfte der Varianz im Meldeverhalten auf zeitlich stabile Faktoren zurückgeht. Entgegen unserer Erwartungen konnten nur 8% der Varianz auf Unterschiede zwischen Lehrpersonen und 7% auf Unterschiede zwischen Schulstunden zurückgeführt werden, während der größte Teil der Varianz (43%) durch Unterschiede zwischen Schüler*innen erklärt wurde. Die Meldehäufigkeit von Schüler*innen scheint demnach relativ stabil zu sein, wobei der Einfluss von Lehrpersonen und Schulstunden zwar eine bedeutende, jedoch vergleichsweise geringe Rolle spielt.

Im Hinblick auf die reziproken Zusammenhänge zwischen Schülermeldungen und Lernermerkmalen zeigt sich, dass Schülermeldungen zu Beginn des Schuljahres einen signifikanten Effekt auf die Entwicklung des Selbstkonzepts haben ($\beta = .07$). Darüber hinaus zeigt sich, dass Schülermeldungen sowohl einen Effekt auf die Entwicklung der kognitiven Lernaktivität ($\beta = .12$) aufweisen, als auch, dass die kognitive Lernaktivität selbst im positiven Zusammenhang mit Veränderungen in der Meldehäufigkeit steht ($\beta = .12$). Schülermeldungen eröffnen demnach wichtige Lerngelegenheiten, die sowohl mit positiven Entwicklungen in der Kompetenzeinschätzung als auch in der kognitiven Lernaktivität einhergehen. Fördermöglichkeiten und die Rolle von Lehrpersonen werden im Vortrag diskutiert.

Wie beeinflussen individuelle und kontextabhängige Determinanten das aufmerksamkeitsbezogene Verhaltensweisen von Schüler:innen im Unterricht?

Patricia Goldberg[1], Wolfgang Wagner[1], Tina Seidel[2], Kathleen Stürmer[1]
[1]Universität Tübingen, [2]Technische Universität München

Theoretischer Hintergrund

Damit Schüler im Unterricht erfolgreich lernen können, müssen sie aufmerksam sein und sich mit den Lerninhalten auseinandersetzen. Lehrkräfte müssen sich dabei auf beobachtbare Hinweise im Verhalten der Schüler:innen stützen, die anzeigen, ob und wie intensiv sich die Schüler:innen mit dem Lerninhalt befassen (Goldberg et al., 2021). Sichtbare Verhaltensweisen können dabei schülergesteuerte Beiträge und Wortmeldungen, aber auch subtilere Hinweise wie dauerhaft abgewandte Blickrichtungen sein. Ergebnisse früherer Studien deuten darauf hin, dass Unterschiede im aufmerksamkeitsbezogenen Verhalten von Schüler:innen auf verschiedene Faktoren zurückzuführen sein können. So zeigten Jurik et al. (2013), dass individuelle Merkmale von Schüler:innen ihr Verhalten im Klassenzimmer beeinflussen und zu individuellen Unterschieden zwischen den Schüler:innen führen. Darüber hinaus scheinen kontextabhängige Faktoren, die für einzelne Klassen und ihre Prozesse typisch sind (z. B. Klassenkultur und Interaktionsdynamik), sowie bestimmte Anforderungen an die Lernaktivitäten im Klassenzimmer einen Einfluss darauf zu haben, welche Art von aufmerksamkeitsbezogenem Verhalten Schüler:innen zeigen (Brühwiler & Blatchford, 2011; Chi & Wylie, 2014).

Bislang ist jedoch nicht klar, ob und wie diese Faktoren miteinander zusammenhängen und inwieweit sie das Verhalten der Schüler:innen im Verlauf einer Unterrichtseinheit bestimmen. Bestehende Forschungsansätze verwenden in der Regel aggregierte Messwerte für das aufmerksamkeitsbezogene Verhalten von Schüler:innen, wodurch situative Aspekte im Verhalten der Schüler:innen nicht adäquat abgebildet werden können (Diener & Larsen, 2009).

In der vorliegenden Studie wird untersucht, wie sich individuelle Merkmale der Schüler:innen, die Klassenzugehörigkeit, die Aktivitäten im Klassenzimmer und die Qualität des Unterrichts auf das aufmerksamkeitsbezogene Verhalten der Schüler:innen in einer Unterrichtssituation auswirken und welche Unterschiede zwischen den Schüler:innen bestehen.

Fragestellung

Wir haben die folgenden Forschungsfragen untersucht:

1) Wie beeinflussen die individuellen Merkmale der Schüler:innen (kognitive Fähigkeiten, Vorwissen, Interesse und Selbstkonzept), die Klassenzugehörigkeit, die Unterrichtsaktivitäten und die Qualität des Unterrichts (Routinen, Ablenkungen, Grad der kognitiven Aktivierung im Unterrichtsgespräch) das aufmerksamkeitsbezogene Verhalten der Schüler:innen (Varianz innerhalb einer Person) in einem Unterrichtssetting?

2) Inwieweit sind die individuellen Merkmale der Schüler:innen (kognitive Fähigkeiten, Vorwissen, Interesse und Selbstkonzept), die Klassenzugehörigkeit, die Unterrichtsaktivitäten oder die Qualität des Unterrichts (Routinen, Ablenkungen, Grad der kognitiven Aktivierung im Unterrichtsdialog) für die Unterschiede zwischen den SchülerInnen in ihrem aufmerksamkeitsbezogenen Verhalten während des Unterrichts verantwortlich?

Methode

Die vorliegende Studie basiert auf N=10 Unterrichtsvideos (Länge jeweils etwa 20 min) aus dem Fach Mathematik, die im Rahmen einer Studie von Seidel et al. (2016) aufgenommen wurden, und die Lehrenden-zentrierten Abschnitte des Unterrichts zeigen. Wir haben das beobachtbare Verhalten von N=199 Schüler:innen auf einer eindimensionalen Skala entlang der Videolänge in Sekundenschritten manuell kodiert (Goldberg et al., 2021). Zudem wurden die Unterrichtsaktivitäten (z.B. Lehrenden-Vortrag, Hefteintrag, Diskussion) von zwei Beobachtenden kodiert (κ = .92). Um mögliche Einflüsse der Unterrichtsqualität berücksichtigen zu können, wurden Routinen und Störungen sowie das Klassengespräch ebenfalls von zwei Beobachtenden bewertet.

Um die zweistufige Struktur unserer Daten zu berücksichtigen (d.h., Verhaltensbewertung jede Sekunde innerhalb der Schüler:innen), wurden im Rahmen des Dynamic Structural Equation Modeling Framework (DSEM) mehrstufige, location-scale, autoregressive Modelle verwendet (Asparouhov et al. (2018).

Ergebnisse und ihre Bedeutung

Unsere Ergebnisse zeigen, dass die Klassenzugehörigkeit 84,47% der Varianz und die Gestaltung des Unterrichts bezogen auf die Unterrichtsaktivitäten 61,16% der Varianz zwischen den Schüler:innen erklären konnte. Der Grad der kognitiven Aktivierung im Unterrichtsgespräch konnte mit 46,46% für die Verwendung von Klärungsfragen und 25,07% für die Anregung zum kritischen Denken die größten Anteile bezogen auf Aspekte der Unterrichtsqualität in der Varianz zwischen den Schüler:innen erklären. Individuelle Charakteristiken der Schüler:innen erklärten hingegen nur 7,70% der Varianz.

Diese Ergebnisse legen nahe, dass aufmerksamkeitsbezogenes Verhalten in erster Linie durch kontextbezogene Faktoren bestimmt wird, was die wichtige Rolle der Lehrkräfte, aber auch die Auswirkungen von Faktoren, die in einzelnen Klassenräumen einzigartig sind, unterstreicht.

Sondereranstaltung

D-21: Sonderveranstaltung: RatSWD

Zeit: Mittwoch, 01.03.2023: 13:15 - 15:00 · *Ort:* R14 R02 B07

Evidenz verbessern mit qualitätsgesicherten Daten akkreditierter Forschungsdatenzentren

Kerstin Schneider[1], Daniel Buck[2], Malte Jansen[3], Karen Meyer[4], Thomas Lösch[5]

[1]Universität Wuppertal, Deutschland; [2]FDZ des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung; [3]FDZ des Instituts zur Qualitätsentwicklung im Bildungswesen; [4]FDZ der statistischen Ämter der Länder; [5]FDZ Bildung am Leibniz-Institut für Bildungsforschung und Bildungsinformation

Wissenschaftliche Evidenz kann Wissen von Scheinwissen, News von Fake News, Denken von Querdenken abgrenzen. Für die Generierung empirisch gesicherter Aussagen ist auch in der empirischen Bildungsforschung die Qualität der zugrundeliegenden Daten zentral. Qualitativ hochwertige Daten stellen die vom Rat für Sozial- und Wirtschaftsdaten (RatSWD) akkreditierten Forschungsdatenzentren (FDZ) zur Verfügung. Forschende können qualitätsgesicherte qualitative und quantitative Daten der FDZ für ihre Fragestellungen auswerten und so die Potentiale von Forschungsdaten ausschöpfen. Seit 2004 setzt sich der RatSWD – ein unabhängiger Beirat der Bundesregierung – für den stetigen Ausbau des Zugangs zu hochwertigen Daten von statistischen Ämtern, Ressortforschung- und Forschungseinrichtungen ein. Inzwischen hat sich ein dezentrales Netzwerk von 42 durch den RatSWD akkreditierten FDZ entwickelt. Seit 2020 arbeiten der RatSWD und die FDZ als Teil des Konsortiums für Sozial-, Verhaltens-, Bildungs und Wirtschaftswissenschaften (KonsortSWD) am Ausbau der Nationalen Forschungsdateninfrastruktur (NFDI) mit. Im Rahmen der Session sollen die Arbeit des RatSWD, des KonsortSWD sowie die Services der FDZ für Forschende vorgestellt werden. Insbesondere sollen ausgewählte Datenbestände für die empirische Bildungsforschung präsentiert werden, anhand derer sich dynamische Entwicklungen im Bereich Bildung auch in Krisenzeiten untersuchen lassen.

Session E

Einzelbeiträge

E-01: Soziologische Fragen

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S06 S00 B29
Chair der Sitzung: Anja Grauenhorst

Formale Höherqualifizierung und soziale Ungleichheit

Anja Grauenhorst, Steffen Schindler
Universität Bamberg, Germany

Hintergrund und Fragestellung

Weiterbildung und lebenslanges Lernen haben in den letzten Jahrzehnten zunehmend an Bedeutung gewonnen. Ihnen kommt zudem eine besondere Relevanz zu, um aktuellen Herausforderungen, wie bspw. Fachkräftemangel in bestimmten Branchen, zu begegnen. Zudem trägt der strukturelle und technologische Wandel dazu bei, dass Weiterbildung im Karriereverlauf für Arbeitskräfte erforderlich wird, um die notwendigen Fähigkeiten für sich wandelnde Anforderungen zu erlangen. Auf der anderen Seite wirft die zunehmende Bedeutung von Weiterbildung auch neue Fragen im Zusammenhang mit sozialer Ungleichheit und Mobilität auf. Da Bildung die zentrale vermittelnde Variable des Zusammenhangs zwischen sozialer Herkunft und sozialer Platzierung ist (Goldthorpe 2014), könnte sozial selektive Beteiligung an Weiterbildung diesen Zusammenhang über den Lebensverlauf verändern.

In diesem Beitrag wird ein bestimmter Aspekt der Weiterbildung untersucht: Formale Höherqualifizierung nach Arbeitsmarkteintritt. Es geht dabei um die Frage, ob nach dem Eintritt in den Arbeitsmarkt die Stärke des Zusammenhangs von sozialer Herkunft und Arbeitslohn auf gesamtgesellschaftlicher Ebene durch sozial selektive Muster des Erwerbs weiterer formaler Bildungsabschlüsse verändert wird und falls ja, auf welche Weise.

Der Beitrag ist eine Replikation der Studie von Virdia und Schindler (2019), die herausgefunden haben, dass formale Höherqualifizierung die Stärke des Zusammenhangs zwischen sozialer Herkunft und sozialem Status nicht verändert. In dieser Studie wurde der soziale Status anhand des Internationalen sozioökonomischen Index des beruflichen Status (ISEI) analysiert.

Dies hat zur Folge, dass Effekte der Weiterbildung lediglich im Zusammenhang mit Berufswechseln identifiziert werden können. Vor dem Hintergrund, dass in Deutschland das Ausmaß beruflicher Mobilität im Lebensverlauf eher gering ist (Hillmert 2011; Mayer, Grunow, Nitsche 2010), könnte das Ergebnis der Studie durch die Wahl des gewählten Indikators der Arbeitsmarkterträge beeinflusst sein. Daher erscheint als alternativer zu analysierender Arbeitsmarktertrag der Arbeitslohn besonders interessant, da dieser im individuellen Karriereverlauf auch innerhalb der gleichen Berufstätigkeit variieren kann. Somit folgen wir dem Vorschlag von Virdia und Schindler (2019) und betrachten den Arbeitslohn anstelle des ISEI.

Theoretischer Rahmen

Virdia und Schindler (2019) zufolge beeinflussen vor allem drei Faktoren das Niveau sozialer Ungleichheit auf gesamtgesellschaftlicher Ebene. Erstens besitzen verschiedene soziale Herkunftsgruppen unterschiedliche Wahrscheinlichkeiten im Karriereverlauf eine formale

Höherqualifizierung nach dem Eintritt in den Arbeitsmarkt durchzuführen. Dies ist vor allem auf Unterschiede in den ursprünglich beim Arbeitsmarkteintritt vorliegenden Bildungsabschlüssen und Arbeitsmarktpositionen zurückzuführen. Zweitens unterscheidet sich aufgrund der Unterschiede in den Ausgangspositionen die Qualität der Höherqualifizierung zwischen den unterschiedlichen, zu analysierenden Gruppen. Hierbei ist die überbrückte Distanz zwischen anfänglich und neu erworbenem Bildungsabschluss und die damit verbundene Einkommensprämie von Bedeutung. Drittens wird das Ausmaß der Auswirkungen der sozial selektiven Beteiligung an Weiterbildung auf gesamtgesellschaftlicher Ebene davon beeinflusst, in welchem Verhältnis die Gruppengrößen der verschiedenen Herkunftsgruppen zueinanderstehen. Vom Zusammenspiel dieser drei Faktoren hängt dann ab, ob Weiterbildung im Lebensverlauf insgesamt zu einer Verringerung oder Verstärkung des Zusammenhangs zwischen der sozialen Herkunft und dem Arbeitslohn führt.

Daten und Methoden

Für die Analysen nutzen wir Umfragedaten des Nationalen Bildungspanels der Startkohorte 6 in Kombination mit administrativen Daten des IAB (NEPS-ADIAB). Mit diesen Daten können wir tagesgenaue Daten zum Arbeitslohn abhängig beschäftigter Personen analysieren. Wir planen Wachstumskurvenmodelle zu berechnen, um die Entwicklung des Zusammenhangs zwischen der sozialen Herkunft und des Arbeitslohns im Karriereverlauf darzustellen und Veränderungen durch formale Weiterbildungsbeteiligung zu identifizieren. Das Brutto-Individualeinkommen ist die zentrale abhängige Variable. Als Indikator für die soziale Herkunft beziehen wir uns auf die EGP-Klasse der Eltern (Erikson 1984).

Ergebnisse

Zu diesem Zeitpunkt liegen noch keine Ergebnisse vor. Bis zur Tagung werden die Analysen jedoch abgeschlossen sein.

Die Bedeutung der zuwanderungsbezogenen Zusammensetzung der Klasse für die mathematischen Kompetenzen geflüchteter Jugendlicher

Hannah Glinka, Oliver Winkler
Martin-Luther-Universität Halle-Wittenberg, Deutschland

Forschungsinteresse

In Folge von Fluchtzuwanderung haben seit 2014 über 1,86 Million Menschen in Deutschland Schutz gefunden, der Großteil wird wahrscheinlich dauerhaft in Deutschland bleiben, weshalb deren Integration notwendig ist. Durch die enge Verknüpfung von Bildung und späteren Erwerbschancen ist insbesondere die Bildungsintegration von Geflüchteten wichtig. In Deutschland ist diese noch unzureichend erforscht. Erste Studien legen nahe, dass ähnliche Disparitäten im Kompetenzerwerb von Geflüchteten auftreten, die von traditionellen Migrant*innengruppen bekannt sind. Der Einfluss von Kompositionseffekten auf die Kompetenzen von Geflüchteten stellt ein Forschungsdesiderat dar, obwohl bekannt ist, dass der Lernkontext den Kompetenzerwerb von Schüler*innen beeinflusst. Vor allem für Deutschland und sein stark stratifiziertes Bildungssystem ist die Betrachtung von Kompositionseffekten interessant, da die Zusammensetzung der Schülerschaft an den Schulformen nicht zufällig, sondern verknüpft mit individuellen sozioökonomischen und zuwanderungsbezogenen

Merkmalen der Schüler*innen ist. Die Forschungsfragen lauten: *Beeinflusst die zuwanderungsbezogene Klassenkomposition die individuellen Mathematikkompetenzen Geflüchteter in der neunten Klasse? Welchen Einfluss haben unterschiedliche Mechanismen?*

Theorie

Theoretisch gehen wir von drei Mechanismen der Zusammensetzung der Schüler*innenschaft aus, die den Kompetenzerwerb von Geflüchteten beeinflussen können:

1) *Sprachlerngelegenheiten*

2) *Lerntempo*

3) *Peereffekte*

Es ist bekannt, dass Geflüchtete häufig in Klassen lernen, die einen hohen Anteil Jugendlicher mit Zuwanderungserfahrung aufweisen, was auf die Verknüpfung von Migrationsstatus, niedriger sozioökonomischer Herkunft und niedrigen Leistungen zurückzuführen ist. In Klassen mit hohem Anteil Jugendlicher mit Zuwanderungserfahrung ist von einem hohen Anteil Jugendlicher nicht-deutscher Herkunftssprache auszugehen, was die Deutsch-Sprachlerngelegenheiten innerhalb des Klassenkontexts für die Geflüchteten verringert, was sich negativ auf ihre individuellen Deutschkenntnisse und dadurch negativ auf ihre Mathematikkompetenzen auswirkt (*H1*). Wir gehen außerdem davon aus, dass Lehrkräfte in Klassen mit hohem Anteil Jugendlicher aus zugewanderten Familien das Lerntempo reduzieren, um den Fokus auf sprachliche Aspekte zu legen, wodurch das Curriculum in Mathematik langsamer unterrichtet wird. Dadurch ist der Mathematikkompetenzerwerb der Geflüchteten reduziert (*H2*). Zuletzt nehmen wir an, dass Peer-Effekte Einfluss auf die Mathematikkompetenzen der Geflüchteten haben können: Niedrigere Leistungsnormen in Klassen mit hohem Anteil Jugendlicher mit Zuwanderungserfahrung wirken sich negativ auf die schulische Motivation aus, was den Kompetenzerwerb der Geflüchteten in Mathematik verringert (*H3a*). Andererseits kann ein hoher Anteil von Jugendlichen aus zugewanderten Familien das Zugehörigkeitsgefühl aufgrund der geteilten Migrationserfahrung innerhalb einer Klasse erhöhen, was sich positiv auf die schulische Motivation und dadurch auf die Mathematikkompetenzen auswirkt (*H3b*). Auch geteilte hohe Aspiration in Klassen mit hohem Anteil Jugendlicher mit Zuwanderungserfahrung können einen positiven Effekt auf die schulische Motivation und damit auf die Mathematikkompetenzen haben (*H3b*).

Methoden

Für unsere Analysen benutzen wir den IQB-Bildungstrend 2018 für die Sekundarstufe. Im finalen Modell werden 490 Geflüchtete miteinbezogen. Wir rechnen eine lineare Regression im Random-Intercept-Mehrebenenmodell (L1: Individualmerkmale, L2: Klassenmerkmale). Dafür sind alle metrischen Variablen standardisiert und Merkmale auf Individualebene gewichtet worden. Die abhängige Variable sind die Mathematikkompetenzen, die als „plausible values“ vorliegen. Kontrolliert wird auf L1 nach Deutschkompetenzen, Schulform, sozioökonomischer Herkunft, Alter und Geschlecht. Auf L2 werden die Mechanismen der Komposition operationalisiert. Die Modelle werden

sukzessiv aufgebaut, sodass ersichtlich wird, welche Kompositionseffekte den Mathematikkompetenzerwerb von Geflüchteten beeinflussen.

Ergebnisse

Deskriptiv zeigt sich, dass Geflüchtete in Klassen unterrichtet werden, die durchschnittlich einen Anteil Jugendlicher mit Zuwanderungserfahrung von 43% aufweisen. Bei Schüler*innen ohne Zuwanderungserfahrung liegt dieser Anteil bei 21%. Auch der Anteil Jugendlicher nicht-deutscher Herkunftssprache und mittlerer HISEI in einer Klasse unterscheiden sich signifikant zwischen Geflüchteten und Schüler*innen ohne Zuwanderungserfahrung. Multivariate Ergebnisse machen sichtbar, dass vor allem die mittleren Deutschkompetenzen in einer Klasse (H2) den Kompetenzerwerb der Geflüchteten positiv beeinflussen. H2 ist damit bestätigt. Es ist ein negativer Einfluss des Anteils Jugendlicher nicht-deutscher Herkunftssprache (H1) auf die Kompetenzen sichtbar, was ebenfalls unseren theoretischen Annahmen entspricht. Der Anteil Jugendlicher aus zugewanderten Familien in einer Klasse hingehen hat einen positiven Effekt auf die Mathematikkompetenzen. Alle anderen Kompositionseffekte verfehlen das Signifikanzniveau.

Sport, basteln oder streaming? Die Freizeitgestaltung von Schüler:innen nach sozialer Herkunft und Geschlecht und der Zusammenhang mit schulischen Merkmalen

Josefine Lühe[1], Dave Möwisch[2]
[1]DIPF, Germany; [2]LIfBi, Leibniz-Institut für Bildungsverläufe

Lernen findet nicht nur in der Schule oder im Rahmen non-formaler Lernangebote wie bspw. Musikkursen, sondern „potentiell immer und überall“ (Rauschenbach 2008) statt. Wenn Kinder ihren Interessen nachgehen, erlangen sie dabei nicht nur Wissen und Können, sondern darüber hinaus auch vielfältige für das Lernen selbst notwendige Voraussetzungen (z.B. Konzentration, Regeleinhaltung, Geschicklichkeit, Zeiteinteilung, Durchhaltevermögen, Kreativität) (Hössl et al. 2002). Wie und womit Kinder und Jugendliche ihre freie Zeit abseits von Schule und auch non-formalen Lernangeboten verbringen, ist damit – und insbesondere vor dem Hintergrund der mit der Corona-Pandemie einhergegangenen eingeschränkten Öffnung von Schulen und anderen Bildungseinrichtungen – von Bedeutsamkeit. Vor diesem Hintergrund untersucht der Beitrag die Freizeitgestaltung von Schüler:innen. Dabei wird folgenden Fragestellungen nachgegangen: (1) Zeigen sich in der Freizeitgestaltung Unterschiede nach Geschlecht und sozialer Herkunft? (2) Welche Freizeittypen lassen sich empirisch zeigen und in welchem Zusammenhang stehen diese mit der Geschlechtszugehörigkeit, der sozialen Herkunft sowie schulischen Merkmalen?

Freizeit bietet Kindern und Jugendlichen einen Raum, in dem sie ihre Persönlichkeit bzw. Identität ausleben und auch erproben können – Freizeitaktivitäten bieten Heranwachsenden eine Möglichkeit, mittels derer sie sowohl sich selbst als auch Anderen kommunizieren können: „This is who I am“ oder „This is what I believe I am meant to do“ (Barber et al. 2008, S. 188). Um die Rolle von Geschlecht und sozialer Herkunft für die Wahl von Freizeitaktivitäten zu beschreiben werden sozialisationstheoretische Ansätze verwendet. In Bezug auf Geschlechterdifferenzen wird das Erwartungs-Wert-Modell (Eccles 1983) herangezogen. Bezüglich der Rolle der sozialen Herkunft für die Wahl von Freizeitaktivitäten wird auf das Bourdieusche Habituskonzept (Bourdieu 1987) zurückgegriffen. Bisherige Forschungsarbeiten zeigen, dass Geschlecht und soziale Herkunft das

Freizeitverhalten beeinflussen (Eccles 2015; Simpkins et al. 2015; mpfs 2021). Darüber können bestimmte Freizeitaktivitäten mit einem höheren Bildungserfolg assoziiert werden (Cabane et al. 2016; Felfe et al. 2016; Bartko und Eccles 2003).

Die Untersuchung basiert auf AID:A-Daten und umfasst 2557 Schüler:innen vom Schuleintritt bis zum Alter von 11 Jahren. In der Erhebung wird erfasst, wie häufig die Schüler:innen verschiedene Freizeitaktivitäten ausführen (z.B. Freunde treffen, draußen spielen, Sport treiben, basteln, Musik machen, lesen, rumhängen). Die soziale Herkunft wird mittels des International Socio-Economic Index (ISEI) abgebildet, welcher Angaben über Informationen über den Beruf, das Einkommen und die Bildung kombiniert. Es wird jeweils der höchste Wert des Haushaltes (HISEI) herangezogen. Die Frage nach herkunfts- und geschlechtsspezifischen Unterschieden in der Freizeitgestaltung wird mittels deskriptiver Analysen beantwortet. Darüber hinaus werden in einem zweiten Schritt latente Klassenanalysen in *Mplus* herangezogen, um verschiedene Freizeittypen abzubilden und deren Zusammenhang mit HISEI, Geschlecht und schulischen Merkmalen darzustellen.

In Bezug auf *Geschlechterunterschiede* in der Freizeitgestaltung zeigen die Analysen, dass Mädchen häufiger als Jungen lesen und Bibliotheken besuchen sowie öfter musisch-kreative Freizeitbeschäftigungen (musizieren, basteln, malen) ausüben. Jungen gehen hingegen häufiger Computer- bzw. Handyspielen nach, spielen öfter draußen und treiben häufiger Sport. Bezüglich der Häufigkeit der Internetnutzung sowie in Bezug auf TV, Streaming und Youtube, Freunde treffen und Rumhängen zeigen sich keine Geschlechterunterschiede. Hinsichtlich *sozial bedingter Unterschiede* in der Freizeitgestaltung verdeutlichen die Analysen, dass Kinder mit niedrigerem HISEI seltener lesen und Bibliotheken besuchen. Im Gegensatz dazu machen sie häufiger von digitalen Angeboten (TV, Streaming, Youtube, Computer- bzw. Handyspielen) Gebrauch, treiben jedoch häufiger Sport als Schüler:innen mit höherem HISEI. Bezüglich der Häufigkeit des Treffens mit Freunden sowie draußen spielen und Internetnutzung finden sich keine sozial bedingten Unterschiede. Zudem werden erste Befunde zu den latenten Klassenanalysen der Freizeitprofile vorgestellt. Vorläufige Ergebnisse zeigen, dass sich mittels latenter Klassenanalyse verschiedene Freizeitprofile abbilden lassen, welche wiederum mit Geschlecht und SES in Zusammenhang stehen.

Die vorgelegten Befunde differenzieller Freizeitaktivitäten unterstreichen die Bedeutsamkeit existierender Ungleichheiten hinsichtlich non-formaler Lernangebote, die ohnehin ausgeprägte existierende Ungleichheiten in den verschiedenen Facetten des Bildungserfolgs verstärken können.

The Immigrant-Native Test Score Gap: A Quantile Regression Perspective

Sebastian E. Wenz
GESIS -- Leibniz-Institut für Sozialwissenschaften, Germany

Usually, researchers in education turn to traditional regression models that focus on conditional means—e.g., linear regression or hierarchical linear models (HLMs)—when assessing differences between immigrants and natives in metric outcomes, such as test scores in educational assessments. These models estimate the conditional mean function, $E(y|x)= x\beta$, and, thus, estimate differences between the mean test score of immigrants and the mean test score of natives. However, such models ignore that differences in means between immigrants and natives—or any other two or more

sociodemographic groups, for that matter—are not necessarily the same as differences in other points (e.g., the 10th percentile, the median, or the 90th percentile) of the outcome's distribution.

In my contribution I show how quantile regression (Koenker & Bassett 1978, Koenker 2005) provides a convenient way of assessing differences between immigrants and natives beyond the mean by estimating the conditional quantile function, Q_tau(y | x)=xβ_tau, and, thus, the difference between quantiles tau, 0<tau<1.

I illustrate the typical use of quantile regression for studying immigrants with an analysis of all 79 countries and regions in OECD's PISA 2018 study. I show that, in many countries, the immigrant native educational achievement gap in quantiles differs remarkably from the mean difference. I discuss potential statistical and substantive explanations--the consequences of processes of (self-)selection, immigration, and integreation--for such differences between linear regression models and quantile regression models and show that these differences also vary across countries.

My analyses show that there are countries in which the *immigrant disadvantage* does *not* differ across quantiles (e.g., USA, GBR, DNK), in which it is *larger* for lower quantiles than for higher quantiles (e.g., DEU, FIN, URY), and there are countries in which it is *smaller* for lower quantiles than for higher quantiles (e.g., ALB, BRA, IDN). In some countries that show an *immigrant advantage*, I observe *no* differences across quantiles (e.g., JOR, SAU), in some countries the advantage is *larger* for lower quantiles than for higher quantiles (e.g., MAC, SGP), and in some it is *smaller* for lower quantiles than for higher quantiles (e.g., ARE, PAN). Last but not least, there are countries in which *both* an immigrant disadvantage and and immigrant advantage can be observed when looking at multiple points of the test-score distributions of immigrant and native students: I find a *disadvantage for lower* quantiles but *advantage for higher* quantiles in LVA, MLT, and MYS. The reversed pattern of a *disadvantage for higher* quantiles but *advantage for lower* quantiles I find in ARG.

I conclude by a brief discussion of the most important limitations and misconceptions around quantile regression: In contrast to popular belief, the traditional conditional quantile regression estimator (Koenker & Bassett 1978, Koenker 2005)—without further assumptions—neither allows inference about particular observations at the quantile under study, nor does it provide estimates that can readily be interpreted as changes in quantiles or differences at quantiles of the unconditional distribution of the outcome. Also, the Quantile Regression Model (QRM) neither requires much less assumptions than the Linear Regression Model (LRM), nor is it any better suited for or in any way superior to the LRM when modeling nonlinear relations (see, e.g., Borgen et al 2021, Wenz 2019).

E-02: Fachdidaktisches Wissen

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S06 S00 B32
Chair der Sitzung: Jannis Zeller

Vergleich des Fachdidaktischen Wissens angehender Physiklehrkräfte aus zwei Large-Scale Studien mit Hilfe des Modells der hierarchischen Komplexität

Jannis Zeller[1], Dustin Schiering[2], Christoph Kulgemeyer[3], Knut Neumann[2], Josef Riese[1], Stefan Sorge[2]
[1]RWTH Aachen University; [2]IPN Kiel; [3]Universität Bremen

Theoretischer Hintergrund und Ausgangslage

Das fachdidaktische Wissen (FDW) bzw. international auch Pedagogical Content Knowledge stellt als Teil des Professionswissens eine zentrale Komponente professioneller Handlungskompetenz von Lehrkräften dar (z. B. Shulman, 1986; Baumert & Kunter, 2006). Dabei wurden Konzeptualisierungen des FDW entwickelt, die sich unter anderem bezüglich der fokussierten Facetten unterscheiden (vgl. Gramzow, 2015), wobei in der internationalen Naturwissenschaftsdidaktik unterschiedliche Sichtweisen auf das Konstrukt in einem Konsens-Modell gebündelt wurden (Hume et al., 2019). Im deutschsprachigen Raum wurden beispielsweise im Projekt KeiLa (Sorge et al., 2019) die Facetten *Schülervorstellungen, Instruktionsstrategien, Assessment und Curriculum* betrachtet, während im ProfiLe-P+ Projekt (Vogelsang et al., 2019) anstelle von Assessment und Curriculum *Experimente* und *Fachdidaktische Konzepte* erfasst wurden. In beiden Projekten wurden dabei Quer- und Längsschnitterhebungen zur Erfassung der Entwicklung des Professionswissens bei Lehramtsstudierenden im Fach Physik durchgeführt, unter anderem mit dem Ziel das FDW datengetrieben kriterienorientiert inhaltlich zu beschreiben (Schiering et al., 2022; Zeller et al., 2022) und so in Evaluationen der Individualität besser gerecht zu werden. Diese Beschreibungen sind dabei aus methodischen Gründen unmittelbar an das zugrundeliegende Testinstrument gebunden und lassen sich daher nicht direkt vergleichen. Somit bleibt zunächst unklar, ob die unterschiedlichen methodischen Ansätze über den Projektkontext hinaus generalisierbare Aussagen ermöglichen. Eine Alternative zum Vergleich unterschiedlicher Ansätze bietet das Modell der hierarchischen Komplexität (Commons et al., 2014), welches für Fachwissen angehender Physiklehrkräfte (Woitkowski & Riese, 2017) sowie von Schüler:innen in Chemie (Bernholt, 2010) bereits erfolgreich zur Konstruktion von Kompetenzniveaumodellen genutzt werden konnte.

Fragestellung

Mit dem Ziel der Herstellung einer projektübergreifenden Vergleichbarkeit und Vereinheitlichung der inhaltlichen Beschreibung von Ausprägungen des FDW auf Basis von empirischen Daten wird in diesem Beitrag ein Testinstrument-unabhängig nutzbares Komplexitätsstufenmodell für FDW-Testitems vorgeschlagen und anhand zweier Large-Scale Studien evaluiert. Dabei wird die folgende Forschungsfrage diskutiert:

FF: Inwieweit lassen sich die Stufen hierarchischen Komplexität des FDW projektübergreifend identifizieren und inhaltlich charakterisieren?

Methode

Aus ProfiLe-P+ liegen Daten zu einem FDW-Testinstrument (Gramzow, 2015) bestehend aus 4 geschlossenem und 20 offenen Aufgaben vor. Das Modell hierarchischer Komplexität wurde mithilfe einer IRT-Modellierung dieses Datensatzes (N=778, Item-Outfits in [0.8, 1.2], EAP-Reliabilität=0.71)

konstruiert. Angelehnt an das Vorgehen von Woitkowski und Riese (2017) wurde ein dreistufiges Komplexitätsmodell mit den Stufen *Fakten* (Reproduktion), *Einstufige Kausalität* (einstufige Argumentationen, Begründungen, Evaluationen, etc.) und *Mehrstufige Kausalität* (mehrstufige Argumentationen etc.) zur Kategorisierung der Aufgaben entwickelt, welches nach kommunikativer Validierung für dieses Testinstrument eine Beurteilerübereinstimmung von κ=0.87 erreichte. Das Komplexitätsstufenmodell wurde anschließend auf das FDW-Testinstrument des KeiLa Projekts bestehend aus 20 geschlossenen und 19 offenen Aufgaben übertragen (N=427, IRT-Modellierung Item-Outfits in [0.8, 1.2], EAP-Reliabilität=0.72). Dabei ergab sich eine Beurteilerübereinstimmung von κ=0.82 nach kommunikativer Validierung. Zur Quantifizierung der Übertragbarkeit wurden deskriptive Maße und lineare Regressionsmodelle verwendet.

Ergebnisse und Schlussfolgerungen

Die Varianz der Aufgabenschwierigkeiten des ProfiLe-P+ Testinstruments konnte in einer linearen Regression der Aufgabenschwierigkeit gegen die Komplexitätszuordnung (F=12.8, p<.001) zu 35 % aufgeklärt werden, was für ein nur dreistufiges Modell als hohe Aufklärung zu deuten ist (Woitkowski & Riese, 2017). Trotz der hohen Einigkeit in der Zuordnung der Aufgaben zu den Komplexitätsstufen auch für das KeiLa Testinstrument wurde ein analoges Modell für dieses nicht signifikant und lieferte eine Varianzaufklärung von nur 5 %. So zeigt beispielsweise ein Faktenitem zur Bedeutung der Basiskonzepte eine deutlich höhere Schwierigkeit als im Modell hierarchischer Komplexität erwartet wäre. Somit ergibt sich, dass die durch die Testinstrumente erfassten Konstrukte zwar verwandt sind, aber das Modell der hierarchischen Komplexität ebenfalls an Grenzen bei der Betrachtung des FDW stoßen kann. Im Vortrag werden anhand weiterer konkreter Testitems der Instrumente im Detail analysiert, wo Gemeinsamkeiten und Unterschiede zu erkennen sind und welche Schwierigkeiten die gemeinsame Modellierung durch ein projektübergreifend feststehendes Komplexitätsmodell mit sich bringt, um somit gemeinsame Schlussfolgerungen über den Kompetenzaufbau angehender Lehrkräfte zu ermöglichen.

Entwicklung fachdidaktischen Wissens zum Einsatz digitaler Medien im Fach Physik

Rike Große-Heilmann[1], Jan-Philipp Burde[2], Josef Riese[1], Thomas Schubatzky[3], David Weiler[2]
[1]Didaktik der Physik und Technik, RWTH Aachen, Deutschland; [2]AG Didaktik der Physik, Eberhard Karls Universität Tübingen, Deutschland; [3]Institut für Fachdidaktik, Universität Innsbruck, Österreich

Theoretischer Hintergrund

Der Einsatz digitaler Medien eröffnet neue Möglichkeiten für den Unterricht. In Physik kann beispielsweise eine neue Qualität der Visualisierung geschaffen werden, um das konzeptionelle Verständnis der Lernenden zu fördern [1,2]. Notwendiges Wissen angehender Lehrkräfte für den Einsatz digitaler Medien wird dem fachdidaktischen Wissen (FDW) zugerechnet, konkret z.B. als Facette *(Digitale) Medien* des FDW in Physik [3]. Diese Facette beinhaltet das Wissen über Möglichkeiten und Anforderungen der inhaltspezifischen Nutzung digitaler Medien im Physikunterricht und ähnelt dem zentralen Wissensbereich *Technological Pedagogical Content Knowledge* (TPCK) im fachübergreifenden TPACK-Modell [4].

Zur Förderung dieses FDW zum Einsatz digitaler Medien sind entsprechende fachdidaktische Lerngelegenheiten notwendig. Das *Synthesis of Qualitative Evidence* (SQD) Modell [5] bietet dazu

einen allgemeinen Ansatz zur methodischen Gestaltung solcher Lerngelegenheiten, der sieben Schlüsselaspekte bezüglich der Vorbereitung angehender Lehrkräfte auf den Einsatz digitaler Medien auf Seminarebene berücksichtigt (z.B. Theorie-Praxis-Verknüpfung, eigenes Gestalten des Medieneinsatzes, Kooperation und Feedback). Zur inhaltlichen Gestaltung der Lerngelegenheiten kann aus fachdidaktischer Sichtweise insbesondere die Nutzung fachspezifischer oder -typischer Medien fokussiert werden [6] (z.B. digitale Messwerterfassungssysteme und Simulationen [7]). Um die Lernwirksamkeit entsprechender universitärer Lerngelegenheiten zu digitalen Medien zu beurteilen, kann die Untersuchung der Entwicklung des FDW zum Einsatz digitaler Medien erste Hinweise liefern.

Fragestellung

In diesem Sinne zielt das vorgestellte Projekt auf die Evaluation physikdidaktischer Seminare zum Einsatz digitaler Medien vierer kooperierender Standorte (Aachen, Innsbruck, Graz und Tübingen) ab, wobei der Fokus in diesem Beitrag auf der Veränderung des FDW zum Einsatz digitaler Medien liegt. Die Seminare folgen einem übergreifenden Lehrkonzept, das sich an den Schlüsselaspekten des SQD-Modells [5] orientiert. Es besteht aus gemeinsam abgestimmten Kerninhalten (lernpsychologische Grundlagen zum Medieneinsatz sowie Designprinzipien, Einsatzmöglichkeiten und empirische Befunde zu physikspezifischen oder -typischen Medien wie digitale Messwerterfassungssysteme, Simulationen und Erklärvideos) sowie standortspezifischen praktischen Anteilen (Schulerprobung in Aachen und Micro-Teachings in Graz und Tübingen). Aufbauend auf der Evaluation hinsichtlich des FDW zum Einsatz digitaler Medien soll durch die Identifikation lernwirksamer Seminarelemente Gestaltungswissen für Lerngelegenheiten zum Einsatz digitaler Medien im Physikunterricht und darüber hinaus gewonnen werden.

Die erste Forschungsfrage adressiert die Untersuchung des Wissenserwerbs in den Seminaren mithilfe eines im Projekt entwickelten Leistungstests zum FDW zum Einsatz digitaler Medien:

FF1: Zeigt sich eine Veränderung im physikdidaktischen Wissen zum Einsatz digitaler Medien über die Seminare der vier beteiligten Standorte?

Die zweite Forschungsfrage fokussiert die Identifikation lernwirksamer Elemente des Lehrkonzepts unter Berücksichtigung der gemeinsamen sowie standortspezifischen Anteile:

FF2: Welche Elemente des gemeinsamen Kerns bzw. der standortspezifischen Teile des Lehrkonzepts sind besonders lernförderlich bzw. eher lernhinderlich?

Methode

Die Messung des FDW über die Seminare (FF1) erfolgt im Pre-Post-Design mittels des im Projekt entwickelten und validierten Testinstruments zum physikdidaktischen Wissen zum Einsatz digitaler Medien (Ergebnisse verschiedener Validierungsstudien in [8]). Bislang konnten N=35 Pre-Post-Datensätze erhoben und mit einem t-Test abhängiger Stichproben analysiert werden. Zur Identifikation wirksamer Seminarelemente (FF2) werden mit einem Teil der getesteten Studierenden retrospektive Leitfaden-Interviews nach der Pre-Post-Befragung durchgeführt. Dabei werden Ursachen für individuelle Veränderungen in den Testantworten erkundet und auf Elemente des

Lehrkonzepts bezogen. Bislang wurden *N*=10 Interviews geführt. Die Leitfaden-Interviews werden nach der qualitativen Inhaltsanalyse nach Kuckartz [9] ausgewertet. Im Studienjahr 2022/23 werden weitere Daten zu FF1 und FF2 erhoben.

Ergebnisse

Die vorläufigen Ergebnisse der Längsschnitterhebungen über die Seminare (FF1) zeigen eine positive Entwicklung des physikdidaktischen Wissens zum Einsatz digitaler Medien ($t(35)=2.553$, $p=0.015$, $d=0.43$). Nach aktuellem Stand der Interviewauswertung (FF2) deutet sich an, dass beispielsweise der eigene Medieneinsatz in der Praxisphase sowie die Auseinandersetzung mit Designprinzipien zum Multimedialen Lernen lernförderlich sind und zu einer Verbesserung der Testantworten führen. Vereinzelt werden manche Distraktoren im Test durch das Kennenlernen der Möglichkeiten einzelner Medien im Seminar scheinbar weniger kritisch bewertet und somit häufiger ausgewählt. Demnach sollte im Seminar verstärkt eigenes Gestalten sowie kritisches Reflektieren von Medieneinsätzen ermöglicht werden.

Die Vorhersage des fachdidaktischen Wissens mit den COACTIV-Daten durch Machine Learning Verfahren

Stefanie Hohenleitner, Stefan Krauss, Sven Hilbert, Martin Fröhlich
Universität Regensburg, Deutschland

Die COACTIV-Studie untersuchte unter anderem den Einfluss verschiedener professioneller Kompetenzen von Mathematik-Lehrkräften und von Aspekten der Unterrichtsqualität auf den Leistungszuwachs der Schüler*innen. Dabei wurde das fachdidaktische Wissen als stärkster Prädiktor identifiziert (Baumert & Kunter, 2011). Diese Ergebnisse werden auch von anderen Studien bestätigt, wie beispielsweise das Projekt der Michigan-Gruppe (Hill et al., 2005). Daraus ergibt sich die Frage, welche Eigenschaften einer Lehrkraft, Aspekte des Unterrichts beziehungsweise weiterer erhobener Kontextvariablen wiederum einen positiven Effekt auf das fachdidaktische Wissen der Lehrkraft und damit indirekt auch auf den Leistungszuwachs der Schüler*innen haben könnten.

Bisher wurden sowohl in der COACTIV-Studie als auch in anderen Untersuchungen in diesbezüglichen Pfad- beziehungsweise Strukturgleichungsmodellen und unabhängig vom avisierten Zielkriterium in der Regel nur *einzelne* beziehungsweise nur eine *beschränkte* Anzahl von Konstrukten als Prädiktoren modelliert. Dazu zählen sowohl spezifische Facetten übergeordneter Kompetenzaspekte wie der Überzeugungen, der motivationalen Orientierung oder der selbstregulativen Fähigkeiten als auch Merkmale von Unterrichtsqualität wie die kognitive Aktivierung, konstruktive Unterstützung oder Klassenführung (Kunter et al., 2013). Aus diesen Befunden jedoch lässt sich jeweils *nicht* ableiten, welche Variablen die *stärksten* Prädiktoren aus einer großen Anzahl von Möglichkeiten darstellen.

Daher soll folgende Fragestellung in den Blick genommen werden: Welche in der COACTIV-Studie erfassten Lehrkraft-, Unterricht- oder Kontextvariablen (insgesamt wurden über 2.000 Variablen erhoben) zeigen die *stärkste* Assoziation mit dem fachdidaktischen Wissen und sollten daher in zukünftiger Forschung im Zusammenhang mit der Entwicklung von fachdidaktischem Wissen betrachtet werden?

Mithilfe von Machine Learning Verfahren wurden alle COACTIV-Skalen in *einem* Modell zur Vorhersage des fachdidaktischen Wissens implementiert. Dies sind insgesamt über 300 Skalen. Viele Machine Learning Verfahren haben – zum Beispiel im Gegensatz zu Strukturgleichungsmodellen, in denen sich schnell Kollinearitätsprobleme einstellen (Bühner, 2021) – den Vorteil, dass eine beinahe unbeschränkte Anzahl an Prädiktoren implementiert werden kann. Um das Problem des Overfitting zu umgehen, setzt Machine Learning zudem auf *Resampling* Verfahren, welche die Genauigkeit der Vorhersage an unbekannten Daten schätzen (Hilbert et al., 2021). Zusätzlich wurde *Interpretable Machine Learning* implementiert, um die Bedeutsamkeit aller Variablen für die Vorhersage des fachdidaktischen Wissens durch *permutation importance* zu bestimmen (Molnar, 2018). Hierzu wurde ein *random forest* trainiert. Die Datengrundlage bilden dabei insgesamt die 194 COACTIV-Klassen, welche sowohl 2003 als auch 2004 an der Untersuchung teilnahmen. Dabei dienten etwa zwei Drittel der Klassen als Trainingsset und etwa ein Drittel als Testset.

In der Tat erwies sich dabei eine Skala als sehr prädiktiv, die bislang in keiner der insgesamt knapp 200 COACTIV-Publikationen im Zusammenhang mit dem fachdidaktischen Wissen diskutiert wurde, nämlich die Skala *Kooperations- und Absprachebereitschaft mit Kolleg*innen*. Im Sinne von Theorien wie beispielsweise zu professionellen Lerngemeinschaften lassen sich diese Befunde sehr gut einordnen und interpretieren (zu ähnlichen Befunden siehe z. B. auch Horn et al., 2020). Weiterhin zählt die Skala *Qualität der Hausaufgaben* zu den starken Prädiktoren, was ebenfalls ein Befund ist, der mit klassischen konfirmatorischen Methoden bislang in diesem großen Datensatz noch nicht identifiziert werden konnte, auch wenn hier natürlich theoretisch eher eine umgekehrte Kausalitätsrichtung vorliegt.

Diese Analysen legen nahe, dass sich auch bei Konstrukten wie dem fachdidaktischen Wissen, die mit den COACTIV-Daten bereits hinreichend analysiert scheinen, mit Machine Learning Verfahren noch neue und bildungstheoretisch relevante Befunde finden lassen. Dies motiviert die Vorhersage anderer Zielkriterien im COACTIV-Datensatz mit diesen Methoden, nicht zuletzt natürlich auch des Lernzuwachses von Schüler*innen. Im Vortrag sollen weiterhin Vor- und Nachteile rein explorativer Machine Learning Verfahren im Hinblick auf Large-Scale-Datensätze der empirischen Bildungsforschung diskutiert werden (siehe hierzu auch Hilbert et al., 2021).

Fächerübergreifende Wissensfacetten als Teil der professionellen digitalen Kompetenz bei Lehramtsstudierenden: Ergebnisse zur Struktur und zu Korrelationen mit Lerngelegenheiten

Sandra Heine, Kristina Gerhard, Daniela Jäger-Biela, Johannes König
Universität zu Köln, Germany

Theoretischer Hintergrund

Lehrkräfte spielen für die Vermittlung digitaler Kompetenzen eine wichtige Rolle. Dafür müssen (angehende) Lehrkräfte aber selbst mit den entsprechenden Kompetenzen ausgestattet werden (Bedir, 2019). Interessanterweise schneiden aber gerade Lehramtsstudierende im Bereich der ICT-Literacy im Vergleich zu anderen Studierenden schlechter ab (Senkbeil et al., 2020).

Dabei haben sich die Anforderungen an das professionelle Wissen (angehender) Lehrkräfte in den letzten Jahren verändert. Die Herausforderungen, die sich durch die Digitalisierung für die Gesellschaft ergeben, nehmen zunehmend Einfluss auf das Bildungssystem und damit auch auf die

Lehrkräfte (Mußmann et al., 2021; Koehler & Mishra, 2009). Mishra und Koehler berücksichtigen dies durch die Erweiterung des Wissensmodells von Shulman (1987). Indem sie eine Technologie-Komponente ergänzten, entstand das vielzitierte TPACK-Modell mit sieben verschiedenen Wissenskomponenten (Mishra & Koehler, 2006).

Da sich die Digitalisierung auf alle Bereiche des gesellschaftlichen Lebens erstreckt, fordert die Kultusministerkonferenz (KMK) in den Curricula eine Balance zwischen fachlichen und fachübergreifenden Inhalten (KMK, 2021). Fachübergreifend im Sinne des TPACK-Modells adressiert dabei die folgenden Wissenskomponenten: general pedagogical knowledge (GPK), technological knowledge (TK) und technological pedagogical knowledge (TPK). Ein Wissenstest zur Erfassung des TK existiert, nach unserem Kenntnisstand, bisher noch nicht. Daher ist auch die Struktur des fachübergreifenden professionellen digitalen Wissens bislang noch nicht empirisch belegt. Für die zielgerichtete Förderung der Kompetenzen (angehender) Lehrkräfte in ihrer Ausbildungszeit ist eine solche Kenntnis jedoch unerlässlich.

Fragestellung

1. Welche Struktur zeigt sich zwischen den drei Wissenskomponenten general pedagogical knowledge (GPK), technological knowledge (TK) und technological pedagogical knowledge (TPK) von Lehramtsstudierenden fächerübergreifend?
2. Sind Unterschiede im erlangten Wissensniveau bei Lehramtsstudierenden in Abhängigkeit von ihren Lerngelegenheiten erkennbar?

Methode

Die Datenbasis stammt aus der Erhebung eines Monitorings zur Lehramtsausbildung an der Universität zu Köln. Im Sommersemester 2022 wurde bei rund 700 Lehramtsstudierenden das GPK, TK und TPK mithilfe standardisierter Tests erfasst. 386 waren zu diesem Zeitpunkt im zweiten Bachelor-Semester (Rücklauf 28,3%), rund 320 im zweiten Master-Semester (Rücklauf rund 39%). Um das GPK zu erfassen, kam das Instrument aus TEDS-M zum Einsatz (Blömeke et al., 2010). Zur Messung des TPK wurde ein standardisiertes Testinstrument verwendet, das von einem interdisziplinären Forschungsteam entwickelt und bereits in vorherigen Studien verschiedentlich Anwendung fand (Gerhard et al., 2020). Zur Erfassung des TK wurde ein gänzlich neu entwickelter Wissenstest erstmals breit eingesetzt. Dessen Konzeption baut auf aktuellen Erkenntnissen und Überlegungen zum professionellen Wissen über digitale Ressourcen von Lehrkräften (Heine, Krepf & König, 2022) auf und unterscheidet zwischen dem Umgang mit und dem Wissen über digitale Ressourcen. Neben der Erhebung des Wissenstands wurden zudem die jeweiligen Lerngelegenheiten zu den unterschiedlichen Bereichen erfragt (Jäger-Biela et al., 2020). Zu den drei Bereichen TK, TPK, GPK erhielten die Studierenden Listen möglicher behandelter Themen, bei denen sie angeben sollten, ob diese bereits Gegenstand ihres Studiums gewesen waren.

Ergebnisse und Bedeutung

Aufgrund der erst kürzlich beendeten Datenerhebung werden aktuell die Datenanalysen zur Beantwortung der Fragestellungen durchgeführt, sodass lediglich vorläufige Einblicke genannt werden können. Bis zur GEBF-Tagung 2023 werden die nötigen Datenanalysen jedoch abgeschlossen sein und präsentiert werden können. Alle drei Tests erweisen sich reliabel bei beiden Teilstichproben Bachelor- vs. Masterstudierende. Letztere schneiden in allen Tests statistisch

signifikant besser ab. Hinsichtlich Fragestellung 1 erwarten wir vor dem Hintergrund der zentralen Annahmen aus dem TPACK-Modell mittlere Zusammenhänge zwischen TPK und GPK einerseits und TK und TPK, während TK und GPK schwächer zusammenhängen sollten. Mit Fragestellung 2 prüfen wir die Hypothese, dass spezifische Lerngelegenheiten (erfasst über im Studium bereits behandelter Themen aus Sicht der befragten Studierenden) im mittleren Zusammenhang mit dem jeweils getesteten Wissen stehen.

E-03: Professionalisierung von Lehrkräften

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S06 S00 B41
Chair der Sitzung: Anna Holstein

Videobasierte Intervention zur Förderung der Professionellen Wahrnehmung von Feedback angehender Lehrkräfte

Anna Holstein[1], Christopher Prilop[2], Kira Weber[3], Marc Kleinknecht[1]
[1]Leuphana Universität, Deutschland; [2]Aarhus University, Dänemark; [3]IPN Universität Kiel, Deutschland

Theoretischer Hintergrund

Die Professionelle Wahrnehmung von Unterricht gilt als zentrale Grundlage für professionelles Handeln von Lehrkräften (z.B. Blömeke et al. 2014; Seidel et al., 2008). Sie wird als Kompetenz verstanden, relevante Interaktionen und Prozesse im Unterrichtsgeschehen zu erkennen (*Noticing*) und theorie- und wissensgeleitet zu interpretieren (*Knowledge-based-reasoning*) (Seidel & Stürmer, 2014, van Es & Sherin, 2008) und kann durch videobasierte Trainings bereits im Lehramtsstudium gefördert werden (Janik et al., 2009). Während für die Wahrnehmung von Klassenführung Studien zur Effektivität videobasierter Interventionen vorliegen (Autor:innen, 2018, 2021; Gold et al., 2013; Stürmer et al., 2013), gibt es bislang kaum Studien, die die professionelle Wahrnehmung von anderen Dimensionen guten Unterrichts, wie beispielsweise Lehrer:innen-Schüler:innen-Feedback, erfassen.

Die Fähigkeit, lernförderliches Feedback zu geben, zeichnet sich durch eine hohe Komplexität aus und kann einen enormen Einfluss auf das Lernen von Schüler:innen nehmen (z.B. Narciss et al., 2020). Feedback kann in einfacher und elaborierter Form geäußert werden und sich auf die Aufgabe, den Lernprozess oder die Selbstregulation beziehen (Carless, 2020; Hattie & Timperley, 2007) oder Informationen zur Feedbackrichtung (*Feed up*, *Feed back, Feed forward*) enthalten (Brookhart, 2008; Hattie & Timperley, 2007). Aktuelle Studien zeigen, dass sowohl Lehrkräfte als auch Lehramtsstudierende Schwierigkeiten haben, lernförderliches Feedback zu geben (z.B. Autor:innen, 2022; Lotz, 2016; Pohlmann-Rother, 2020). In der vorliegenden Interventionsstudie wird daher untersucht, wie sich ein videobasiertes Training auf die professionelle Wahrnehmung von Feedback von Lehramtsstudierenden auswirkt.

Forschungsfragen

1. Inwieweit verändert sich das *Noticing* von Feedback der Studierenden, die ein videobasiertes Feedback Training erhalten (IG), im Vergleich zu Studierenden, die ein videobasiertes Training zur Klassenführung erhalten (KG)?

2. Inwieweit verändert sich das *Knowledge-based reasoning* von Feedback der Studierenden der IG im Vergleich zur KG?

Methode

Die Interventionsstudie fand im vierten Mastersemester zweier Lehramtsseminare statt. Die Interventionsgruppe (IG, n = 27; 83% weiblich; M_{Alter} = 26.1) und die Kontrollgruppe (KG, n = 20; 75% weiblich, M_{Alter} = 25.29) nahmen an sieben Seminarsitzungen à 90 Minuten mit angeleiteten Analysen fremder Unterrichtsvideos teil. Die IG zum Thema Feedback und die KG zu Klassenführung. Vor der ersten und nach der letzten Sitzung nahmen alle Studierenden an einem videobasierten Test teil, in welchem sie das Feedback einer Lehrkraft schriftlich analysieren sollten.

Diese schriftlichen Analysen wurden anhand eines Kodiermanuals basierend auf dem ITF- (Narciss, 2017) sowie dem Feedbackmodell von Hattie & Timperley (2007) und angelehnt an Autor:innen (2020) in Bezug auf *Noticing* (*Appropriateness*) und *Knowledge-based-reasoning* (*Beschreiben*, *Erklären* und *Entwicklung von Handlungsalternativen*) ausgewertet. Die Interraterreliabilität aller Kategoriensysteme war sehr gut (Kappa Werte zwischen 0,65 und 0,84).

Ergebnisse und Diskussion

Die Ergebnisse der ANOVAs zeigen, dass die IG sich im Vergleich zur KG im Bereich *Noticing* signifikant verbesserte (p = 0,001, d = 0,7). So war die IG in der Lage, mehr relevante Ereignisse in Bezug auf Feedback zu identifizieren, was als wichtige Grundlage für professionelles Handeln gilt (König et al., 2014). Im Bereich *Knowledge-based reasoning* weisen die Ergebnisse daraufhin, dass die IG sich im Vergleich zur KG signifikant in Bezug auf das *Beschreiben* von lernrelevanten Ereignissen verbessern konnte (p = 0,001, d = 0,75). Studien zu Klassenführung zeigten bereits, dass Noviz:innen in ihren Beschreibungen häufig übergeneralisieren (Stürmer et al. 2015) und dass der Einsatz von Videos zu einer höheren Aktivierung deklarativer Wissensbestände führen kann (Autor:innen, 2020) und Studierende dadurch in der Lage sind, Lernsituationen differenzierter zu beschreiben. Unsere Studie kann ähnliche Erkenntnisse für Lehrer:innen-Schüler:innen-Feedback liefern. Die Facetten *Erklären* und *Entwicklung von Handlungsalternativen* zeigten keine signifikanten Unterschiede zwischen IG und KG. Dies kann mit der Komplexität von Feedback zusammenhängen. Im Vergleich zu Klassenführung brauchen Noviz:innen für den Bereich Feedback eventuell spezifischere Prompts während eines videobasierten Trainings. Die Ergebnisse werden auf der Tagung weiter dargestellt und diskutiert.

Kommt es auf das Format an? - Eine empirische Untersuchung von Online- und Präsenzlehrkräftefortbildungen aus einer flächendeckenden Evaluation

Nora Fröhlich[1], Lucia Teuscher[1], Alexandra Dehmel[1], Benjamin Fauth[1,2]
[1]Institut für Bildungsanalysen Baden-Württemberg; [2]Eberhard Karls Universität Tübingen

Onlinelehrkräftefortbildungen bieten im Vergleich zu Präsenzfortbildungen einige Vorteile. Dazu gehören Einsparungen bei den Reisekosten, der Zugang zu Online-Ressourcen sowie die Möglichkeit zur berufsbegleitenden Unterstützung (Dede et al. 2009, S. 9). Doch wie steht es um die Qualität von Onlineformaten, verglichen mit Präsenzfortbildungen? Bisherige Überblicksarbeiten fokussierten sich bei der Frage nach der Qualität von Fortbildungen überwiegend auf Präsenzfortbildungen und wie diese gestaltet sind bzw. wie wirksam diese sind (z. B. Darling-Hammond et al. 2017; Desimone 2009; Lipowsky und Rzejak 2017, 2021). Immer mehr Überblicksarbeiten beschäftigen sich nun auch mit Online-Fortbildungen und zeigen, dass die von Präsenzfortbildungen abgeleiteten Gestaltungsmerkmale auf Online-Formate übertragbar sind (vgl. Bragg et al., 2021; Capparozza et al., in Druck; Surrette & Johnson, 2015). Ob eine Fortbildung wirksam ist, wird neben der Gestaltung der Fortbildung von einer Vielzahl von Faktoren beeinflusst. Unter anderem von den Voraussetzungen der teilnehmenden Lehrkräfte, von Merkmalen der Fortbildenden und davon, wie die Lehrkräfte die Angebote einer Fortbildung wahrnehmen und nutzen. Lipowsky und Rzejak (2019) fassen diese Aspekte in einem Angebot-Nutzungs-Modell zusammen, anhand dessen Fortbildungen auch evaluiert werden können. Lipowsky & Rzejak (2021) wenden dieses Modell auch auf Online-

Fortbildungen an. Allerdings scheint es Unterschiede in Bezug auf bestimmte Fortbildungsmerkmale zu geben. Erste Studien weisen darauf hin, dass der Kollaboration und der Interaktivität zwischen den Teilnehmenden in Online-Formaten eine wichtigere Rolle zukommt, da der informelle Kontakt schwerer herstellbar ist und Körpersprache sowie Gesichtsausdrücke weniger stark zum Ausdruck kommen können (Elliott, 2017; Quinn et al., 2019). Ein zentraler Gelingensfaktor in Online-Formaten ist deshalb, dass die „Social Presence" (Garrison, 2017) fokussiert wird (Capparozza et al., in Druck). Die Social Presence beschreibt soziale Prozesse während der Fortbildung, wie bspw. den Aufbau einer Beziehung zueinander sowie die offene Kommunikation.

Für diesen Beitrag werden mögliche Unterschiede genauer untersucht: Gibt es bei der Einschätzung der offenen Kommunikation, des Austauschs, der Atmosphäre, der Qualität der Fortbildungsunterlagen sowie der kognitiven Aktivierung Unterschiede zwischen Online- und Präsenzfortbildungen? Womithängen mögliche Unterschiede zusammen? Zur Beantwortung dieser Fragen werden die Daten von ca. 3.000 Fortbildungsveranstaltungen aus dem Schuljahr 2021/22 ausgewertet (Online- und Präsenzfortbildungen). Diese Daten wurden im Rahmen einer flächendeckenden Evaluation der öffentlich angebotenen Fortbildungen in Baden-Württemberg gewonnen. Der eingesetzte Fragebogen orientiert sich an dem Angebot-Nutzungs-Modell von Lipowsky & Rzejak (2019). Im Fokus steht hierbei die erste Ebene des Fortbildungserfolgs im Modell, welche über die unmittelbare Rückmeldung, Zufriedenheit und Akzeptanz der Lehrkräfte erfasst wird.

In einer ersten Voruntersuchung der Daten eines vorläufigen Datensatzes (Erhebungszeitraum November 2021 bis Februar 2022, 914 Veranstaltungen) deuten die deskriptiven Ergebnisse an, dass die Skala „Offene Kommunikation" (adaptiert von Ammenwerth, 2017) sowie die Items „Atmosphäre" (Vigerske, 2017) und „Austausch" (ebd.) auf der Fortbildungsebene in Onlineformaten negativer bewertet werden als in Präsenzformaten. Dahingegen wird die Qualität der Fortbildungsmaterialien (Vigerske, 2017) in Onlinefortbildungen positiver eingeschätzt. Tiefergehende Analysen wurden bisher noch nicht durchgeführt, da der finale Datensatz für das Schuljahr 2021/22 erst Ende November vorliegen wird. Geplant sind Mehrebenenmodelle zur Untersuchung der in den Vorabanalysen gefundenen Unterschieden sowie zur Untersuchung möglicher Unterschiede in der Einschätzung der kognitiven Aktivierung (adaptiert von Ramm et al., 2006). Berücksichtigt werden dabei verschiedene Kovariaten: z.B. fachspezifischer Fokus (Eigenentwicklung) und Zufriedenheit (adaptiert von Vigerske, 2017), sodass auf der GEBF 2023 Ergebnisse vorgestellt werden können.

Da die Daten aus einer flächendeckenden Evaluation stammen, umfassen sie eine große Anzahl an Fortbildungen. Diese wurden i.d.R. nicht wissenschaftlich begleitet und bieten daher einen Einblick in die Fortbildungspraxis unter „Realbedingungen". Dadurch kann ein spannendes und relevantes Feld in der Fortbildungsforschung nun systematisch untersucht werden. Der Datenumfang ermöglicht darüber hinaus differenzierte Analysen über verschiedene Fortbildungen hinweg. Gleichzeitig können praxisrelevante Erkenntnisse gewonnen werden.

Reflexionskompetenz von Lehramtsstudierenden – inklusiven Unterricht wahrnehmen, analysieren und kritisch bewerten als Professionalisierungsaufgabe

Daria Ferencik-Lehmkuhl[1], René Schroeder[2], Silvia Fränkel[1]
[1]Universität zu Köln, Deutschland; [2]Universität Bielefeld, Deutschland

Reflexionskompetenz wird als Kernaspekt professioneller Handlungskompetenz von Lehrkräften betrachtet (Körkkö, Kyrö-Ämmälä & Turunen, 2016; Leonhard & Rihm, 2011). Die Förderung der Reflexionskompetenz sollte deshalb zu Beginn des Studiums angebahnt werden. Um geeignete Angebote machen zu können, muss zunächst der Ist-Stand von Studienanfänger:innen erhoben werden. Während sich einige Studien mit der Reflexionskompetenz von Studierenden in fortgeschrittenen Semestern (Abels, 2011) oder ausgebildeten Lehrkräften (z. B. Abels, 2019) auseinandersetzen, ist bisher unzureichend untersucht, wie sich die Ausgangslage darstellt, also die Reflexionskompetenz bei Studienanfänger:innen ausgeprägt ist (Leonard & Rihm, 2011). Zudem existieren nur wenige Studien, die sich explizit mit der Reflexionskompetenz in Bezug auf inklusiven Unterricht auseinandersetzen (Abels, 2019; Häcker & Walm, 2016).

In diesem Beitrag wird Reflexion als eine im Studium einzuübende Kompetenz definiert, welche u. a. durch Professionswissen, berufsbezogene Überzeugungen, motivationale Orientierungen und selbstregulative Fähigkeiten beeinflusst wird (Baumert & Kunter, 2006). Reflexionen setzen dabei eine gewisse Distanz zum Erlebten bzw. Beobachteten voraus (Häcker, 2019). Reflexionskompetenz erfordert eine Verzahnung von theoretischem Wissen mit Handlungswissen und befähigt zu einer kritischen Auseinandersetzung mit professionellem Handeln. Gerade im Kontext inklusiver Bildung wird die Komplexität im System erhöht (Forlin, Keen & Barrett, 2008). Dabei geht inklusiver Unterricht mit diversen Zielkonflikten wie differenter Leistungsbewertung oder der (Neu-)Aushandlung von Autonomie in Kooperation (Neumann, 2019) einher, wodurch eine stetige Reflexion unterrichtlichen Handelns umso notwendiger erscheint. Ausgehend von einem dialektischen Inklusionsverständnis (Lindmeier & Lütje-Klose, 2015), welches grundsätzlich alle Schüler:innen, aber insbesondere vulnerable Gruppen berücksichtigt, wird die Fähigkeit zur Reflexion inklusiven Unterrichts als notwendige Aufgabe aller Lehrämter und somit als Querschnittsaufgabe der Lehrkräftebildung verstanden.

Vor diesem Hintergrund wurde an der Universität zu Köln (UzK) der Zertifikatskurs „Handlungswissen Inklusion“ im Rahmen des QLB-geförderten Projekts „Zukunftsstrategie Lehrerbildung" (ZuS) entwickelt. Hierfür wurden Aktivitäten der verschiedenen Bereiche und Projekte im Themenfeld Inklusion an der UzK gebündelt und als fächerübergreifendes Zertifikat (Dauer: 5 Semester) angeboten. Lehramtsstudierenden soll zu Anfang ihres Studiums die Chance geboten werden, in verschiedenen Workshops sowie weiteren Veranstaltungen inklusionsspezifische und praxisrelevante Themen zu fokussieren (Fränkel et al., 2022). Der Auf- und Ausbau der Reflexionskompetenz ist Schwerpunkt des Zertifikats, welches durch ein E-Portfolio begleitet wird. Der Zertifikatskurs fand – zunächst als Pilotprojekt – ab dem Sommersemester 2021 an der UzK mit 46 ausgewählten Bachelor-Studierenden statt.

Die Studie wurde im Rahmen des HWI-Projektes durchgeführt. Die vorzustellende Teilstudie erhebt die Reflexionskompetenz im Sinne der Reflexionstiefe (Hatton & Smith, 1995; Leonhard & Rihm, 2011) der Studierenden zu Beginn ihres Studiums bzw. des Zertifikatskurses. Das auf theoretischer Ebene genutzte Reflexionsmodell (Hatton & Smith, 1995) ist kompetenztheoretisch ausgerichtet und beschreibt vier Niveaustufen der Reflexion, welche innerhalb eines für den inklusiven Unterricht adaptieren Kategoriensystems (Fränkel, Ferencik-Lehmkuhl & Schroeder, im Druck) operationalisiert

wurde. Die zentrale Fragestellung der Studie lautet: Wie ist die Reflexionskompetenz entlang der Reflexionsstufen von Lehramtsstudierenden zu Beginn ihres Studiums in Bezug auf inklusiven Unterricht ausgeprägt? Die Reflexionskompetenz der Studienanfänger:innen wurde mittels videovignettenbasierter Interviews (n = 8) erhoben und mit dem adaptierten Kategoriensystem im Rahmen einer qualitativen Inhaltsanalyse (Mayring, 2010) ausgewertet. Für die genutzten Indikatoren konnten gute Intercoder-Reliabilitäten festgestellt werden (κn = 0.85).

Die Ergebnisse zeigen unter anderem, dass die Studierenden zu Beginn ihres BA-Studiums bereits auf einem höheren Niveau reflektieren können, es ihnen aber noch schwerfällt, die höchste Stufe der Reflexion (Critical Reflection) zu erreichen und Handlungsalternativen abzuwägen. Mögliche Gründe könnten begrenztes Theorie- bzw. Kontextwissen zu Inklusion sein (Faix, 2020; Lorenz et al., 2020). Die Chance und gleichzeitig Herausforderung universitärer Lehrkräftebildung ist es, studentische Überzeugungen zum Lehren und Lernen zu diskutieren, mit theoretischem Wissen anzureichern und inklusionsorientierte handlungsleitende Werte und Überzeugungen aufzubauen (Faix, 2020). Im Rahmen des Vortrags wird die Studie ausführlicher vorgestellt sowie die Tragweite der Ergebnisse im Kontext der ersten Phase der Lehrkräftebildung diskutiert.

E-04: Entwicklung von Testinstrumenten

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S06 S02 B06
Chair der Sitzung: Marc Rodemer

Messung Kognitiver Belastung in Lernvideos mit Pupillometrie

Marc Rodemer[1], Jessica Karch[2], Sascha Bernholt[3]
[1]Universität Duisburg-Essen, Deutschland; [2]Tufts University, Boston, MA, USA; [3]Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik, Deutschland

Einleitung

Effektiv gestaltete Lernmaterialien sind unverzichtbar für das Erlernen komplexer Lerninhalte. Die aufgewendete kognitive Belastung beim Lernen wird entweder durch die Schwierigkeit der Aufgabe (intrinsische kognitive Belastung (ICL)) oder durch ihr Design (extrinsische kognitive Belastung (ECL)) induziert (Sweller et al., 1998, 2019). Ein Ziel von Instruktionsdesign ist die Minimierung von ECL, um genügend Arbeitsgedächtnisressourcen für das Verständnis des Lernmaterials freizuhalten. Kognitive Belastung kann anhand physiologischer Maße oder Selbstberichte erfasst werden. Ein Indikator für die kognitive Belastung ist die Pupillenerweiterung, die mit einem Eyetracker gemessen werden kann (für Reviews siehe Beatty & Lucerno-Wagoner, 2000; Just et al., 2003; van der Wel & van Steenbergen, 2018), jedoch haben wenige Arbeiten den Einfluss unterschiedlicher Belastungsarten auf die Pupillenerweiterung untersucht.

Um den Zusammenhang zwischen Pupillenerweiterung und kognitiver Belastung zu untersuchen, haben wir im Instruktionsdesign auf Basis multimedialer Lernprinzipien verschiedene Hervorhebungstechniken variiert (de Koning et al., 2009; Mayer, 2014; van Gog, 2014). Hervorhebungen verringern bekanntermaßen die kognitive Belastung, indem die Aufmerksamkeit der Lernenden auf relevante Teile des Lernmaterials gelenkt wird. Hierbei haben wir drei Hervorhebungsbedingungen verglichen: sequentielle (dynamisch), permanente (statisch) und keine Hervorhebung (Kontrolle). Um die Wirkung dieser Designentscheidungen zu untersuchen, wurden etablierte Fragebögen zur Selbsteinschätzung (Klepsch et al., 2017) und Pupillendurchmesser mit einem Eyetracker gemessen.

Forschungsfragen

FF1: Welche kognitive Belastungsart wird reduziert, wenn Lernvideos mit entweder dynamischen, statischen oder keinen Hervorhebungen angesehen werden?

Basierend auf der Literatur erwarten wir eine Reduktion von ECL in absteigender Reihenfolge von Kontrolle, statisch und dynamisch. Darüber hinaus nehmen wir an, dass ICL aufgrund des Aufgabendesigns konstant bleibt.

Darüber hinaus ist die Beziehung zwischen Pupillenerweiterung und Belastungsarten weitestgehend unklar. Daher lautet die zweite Forschungsfrage:

FF2: Zeigt die Pupillometrie Unterschiede im Pupillendurchmesser an, wenn die kognitive Belastung durch die experimentellen Bedingungen variiert wird?

Unsere Hypothese ist, dass eine entsprechende Abnahme der Pupillenerweiterung zu beobachten ist, wenn ECL über die drei Hervorhebungsbedingungen hinweg abnimmt.

Methodik

An dieser Studie nahmen 28 Studierenden in einem randomisiert-kontrollierten Within-Design teil. Wir haben Lernvideos zu Reaktionsmechanismen der organischen Chemie entwickelt, weil (1) kleine Änderungen an Videos vorgenommen werden können, während die Gesamtgestaltung konstant bleibt, und (2) Reaktionsmechanismen bekanntermaßen visuell und konzeptionell anspruchsvoll sind (für Reviews siehe Daniel et al., 2018; Graulich, 2015). Pupillendurchmesser wurden während der Videobetrachtung mit einem Tobii Pro-Spectrum gemessen.

Ergebnisse

Die Ergebnisse zeigen erwartungsgemäß keinen Unterschied in den Selbstberichten zu ICL zwischen Hervorhebungsbedingungen ($F(2,83)=0.26$, $p=.769$, $\eta^2=.05$). Im Fall von ECL wurde ein signifikanter Haupteffekt für die Hervorhebungsbedingungen gefunden ($F(2,83)=8.89$, $p<.001$, $\eta^2=.19$). Paarvergleiche zeigen einen signifikant niedrigeren ECL für dynamische Hervorhebungen im Vergleich zu statischen oder keinen Hervorhebungen (beide $p<.001$). Es gab keinen Unterschied zwischen der statischen und der Kontrollbedingung ($p=1.00$).

Hinsichtlich der Pupillenerweiterung zeigte sich ein signifikanter Haupteffekt für die Hervorhebungsbedingung ($F(2,83)=3.24$, $p=.045$, $\eta^2=.08$). Paarvergleiche zeigen eine größere Pupillenerweiterung für dynamische Hervorhebungen im Vergleich zu der Kontrollbedingung (p=.001). Es gab keinen signifikanten Unterschied zwischen der statischen und der Kontrollbedingung (p=.065) und zwischen der statischen und der dynamischen Bedingung (p=.195).

Korrelationskoeffizienten zwischen selbstberichteten kognitiven Belastungsskalen (ICL und ECL) und Pupillenerweiterung weisen auf eine negative Assoziation zwischen ECL und Pupillenerweiterung hin ($r(55)=-0.25$, 95%*CI*[-0,48, -0,02], $p=.06$), d. h. wenn eine höherer ECL angeben wurde ist die Pupillenerweiterung negativer, was auf eine kleinere Pupillengröße hinweist. Der Zusammenhang zwischen Pupillenerweiterung und ICL ist ebenfalls negativ, aber kleiner und nicht signifikant ($r(55)=-0.10$, 95%*CI*[-0,30, 0,17], $p=.46$).

Diskussion

Unsere Ergebnisse weisen auf eine stärkere Berücksichtigung von Instruktionsdesign zur Reduzierung der kognitiven Belastung hin. Es wird diskutiert, dass eine Zunahme der Pupillengröße im Vergleich zum Ausgangswert mit mentaler Anstrengung korreliert, die durch eine Reduzierung von ECL ermöglicht wird (Hess & Polt, 1964; Kahneman & Beatty, 1966). Basierend auf dieser Annahme ist mehr systematische Forschung erforderlich, die verschiedene Belastungsarten in Instruktion untersucht.

Erfassung des kollaborativen diagnostischen Problemlöseprozesses unter Auszubildenden zum*r Kfz-Mechatroniker*in

Dave Rexhäuser[1], Anika Radkowitsch[2], Constanze Richters[3], Louise Kaseler[1], Peter Hesse[1], Inga Glogger-Frey[4], Julius Meier[4,5], Stephan Abele[1]

[1]Technische Universität Dresden; [2]IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel; [3]Ludwig-Maximilians-Universität München; [4]Universität Erfurt; [5]Albert-Ludwigs-Universität Freiburg

Kollaboratives Problemlösen gewinnt in der Praxis technischer Berufe wie dem des*r Kfz-Mechatroniker*in zunehmend an Bedeutung. Gründe hierfür liegen insbesondere in steigenden technischen Komplexitäten und zunehmender Wissenszentralisierung bei den Fahrzeugherstellern.

Trotz hoher berufspraktischer Relevanz spielt kollaboratives diagnostisches Problemlösen in Lehrplänen der beruflichen Bildung keine wesentliche Rolle, was u.a. auf fehlende theoretische Konzeptionen und empirisch geprüfte Messverfahren zurückzuführen sein dürfte (Kultusministerkonferenz, 2013). Anhand von Modellen zur Beschreibung des Diagnoseprozesses in der Kfz-Störungsdiagnostik (Meier et al., 2022) und zur Beschreibung der kollaborativen Diagnostik im medizinischen Kontext (Radkowitsch et al., 2020) kann dieser Prozess auf 2 Ebenen beschrieben werden. Dabei werden diagnostische Aktivitäten (Generieren von Hypothesen, Generieren von Evidenzen, Interpretation von Evidenzen) sowie kollaborative Aktivitäten (Teilen von Informationen, Elizitieren von Informationen, Koordination der Kollaboration, Verhandlung, Aufrechterhaltung der Kommunikation) unterschieden.

Im Beitrag präsentieren wir Ergebnisse zur Überprüfung des theoretischen Ansatzes anhand einer simulationsbasierten Erfassung des diagnostischen Problemlöseprozesses unter Auszubildenden der Kfz-Mechatronik. Weiterhin prüfen wir Zusammenhänge zwischen Merkmalen des Prozesses und der Genauigkeit der Diagnose (Diagnoseakkuratheit). Folgende Forschungsfragen werden untersucht: (1) Welche diagnostischen und kollaborativen Aktivitäten können im Prozess gefunden werden? (2) Welcher Zusammenhang besteht zwischen Merkmalen des Prozesses und der Diagnoseakkuratheit?

Insgesamt lösten 128 Auszubildende ($N_{weiblich}$=7; Alter: M=20.72 Jahre; SD=3.01) der Kfz-Mechatronik in Dyaden (N=64) je zwei Kfz-Störungen in einer validierten Computersimulation (Gschwendtner et al., 2009). Innerhalb jeder Dyade übernahm je ein*e Auszubildende*r die Rolle eines*r Mitarbeitende*n in einer Kfz-Werkstatt bzw. einer Kfz-Servicehotline. In beiden Rollen standen jeweils unterschiedliche Funktionen und Informationen zur Verfügung, was eine Zusammenarbeit notwendig machte. Während der Fallbearbeitung kommunizierten die Dyaden über einen Text-Chat. Chatinhalte wurden aufgezeichnet und anschließend anhand eines aus der Theorie entwickelten Codierungsschemas ausgewertet. Bei der Codierung diagnostischer (κ=.89) und kollaborativer Aktivitäten (κ=.77) wurden gute Interrater-Reliabilitäten erzielt. Die Qualität des kollaborativen Diagnoseprozesses wurde mittels einer durch Kfz-Experten erstellten Musterlösung bewertet.

(1) In den Protokollen konnten die theoretisch angenommenen Aktivitäten gefunden werden (Tabelle 1). Der größte Anteil der diagnostischen Aktivitäten entfiel auf das: „Generieren von Evidenzen“ (54%) sowie „sonstige Aspekte“ (z.B. Small Talk; 34%). Die „Interpretation von Evidenzen“ (8%) und das „Generieren von Hypothesen“ (4%) konnte seltener beobachtet werden. Innerhalb der kollaborativen Aktivitäten nahmen die Aktivitäten „Teilen von Informationen“ (34%), „Koordination der Kollaboration“ (28%) und „Aufrechterhaltung der Kommunikation“ (21%) einen großen Anteil in Anspruch. „Elizitieren von Informationen“ (12%) sowie „Verhandeln“ (5%) wurde selten gefunden. Im Prozessverlauf konnten Unterschiede im Zusammenhang der Aktivitäten festgestellt werden. Das Teilen und Elizitieren von Informationen ist über den gesamten diagnostischen Prozess mit einem hohen Anteil (61%-71%) vertreten, ausgenommen in der Kategorie „sonstige Aspekte“, die durch Aufrechterhaltung der Kommunikation (62%) sowie Koordination (35%) bestimmt wird. Koordination tritt vermehrt beim Generieren von Evidenzen (26%), aber weniger beim Generieren von Hypothesen (12%) und der Interpretation von Evidenzen (16%) auf. Mit der Verhandlung verhält es sich entgegengesetzt: 2% beim Generieren von Hypothesen gegenüber 20% beim Generieren von Evidenzen bzw. 23% beim

Interpretieren von Evidenzen. (2) Die Analysen zeigen keine signifikanten Zusammenhänge zwischen der Häufigkeit des Auftretens der Aktivitäten und der Diagnoseakkuratheit. Detailanalysen der Dyaden-Protokolle konnten jedoch signifikant positive Zusammenhänge zwischen der Diagnoseakkuratheit und dem korrekten Austausch besonders relevanter Evidenzen (ρ=.355; p=<.006) zeigen. Zur Identifizierung besonders relevanter Evidenzen dienten Expertenlösungen.

Die Ergebnisse legen nahe, dass das eingesetzte Messverfahren verlässlich Differenzierungen im Prozess zwischen den Dyaden aufdecken kann und sich der gewählte Modellansatz zur Beschreibung des Prozesses kollaborativ diagnostischen Problemlösens in der Kfz-Störungsdiagnose eignet. Dementsprechend gehen wir davon aus, dass das Messverfahren zur Evaluation des Effektes einschlägiger Interventionen geeignet ist. Zudem zeigte sich, dass es bei der Kollaboration hinsichtlich der Diagnoseakkuratheit nicht auf die Aktivitätshäufigkeiten, sondern darauf ankommt, diagnoserelevante Evidenzen korrekt zu elizitieren und zu teilen. Dies spricht dafür, in Interventionen zur kollaborativen Kfz-Störungsdiagnostik genau diese Aktivitäten zu fördern.

Messung der Fähigkeit zum Lösen komplexer domänenorientierter Probleme mit komplexen Itemsequenzen

Thilo Ketschau
Christian-Albrechts-Universität Kiel, Germany

Der Umgang mit Unsicherheit ist auch wesentlicher Bestandteil der Auseinandersetzung mit (komplexen) Problemen, kann doch nach Dörner (1987) die Unklarheit über Handlungsmöglichkeiten und Zielkriterien eine Anforderungssituation subjektiv als Problem wahrnehmbar machen.

Der vorliegende Beitrag stellt eine Studie vor, deren Ziel die Konstruktion eines auf geschlossenen Itemformaten basierenden Instruments zur Messung komplexer Problemlösefähigkeit war, welches zur Sicherstellung einer hohen Auswerteeffizienz auf offene bzw. qualitative Elemente verzichtet. Die für das Instrument konstruierten Problemsituationen wurde mit starken Bezügen zum Prinzip Nachhaltigkeit (vgl. United Nations World Commission on Environment and Development 1987) in die Domäne des nachhaltigen Wirtschaftens (Schaltegger & Wagner 2006) eingebettet, um die Diagnose der Fähigkeit zum Lösen komplexer domänenorientierter Probleme unter Studierenden unterschiedlicher Studiengänge zu ermöglichen. Die dem Instrument zugrunde liegenden Problemsituationen wurden dabei als domänenorientierte Probleme angelegt, d.h. entgegen domänenspezifischen Problemen wurde zunächst formal der Problemraum konstruiert und dieses formale Konstrukt anschließend semantisch in den Kontext nachhaltigen Wirtschaftens eingebettet.

Die Operationalisierung der adressierten Fähigkeit zum Lösen komplexer Probleme basiert auf den durch Dörner (1986) formulierten Systemmerkmalen Komplexität, Vernetztheit, Intransparenz, Polytelie und Dynamik. Der im weiteren Vorgehen gewählte testtheoretische Zugang orientierte sich in der Folge eher an den klassischen Maximal Complex Systems, bspw. Lohhausen (Dörner 1989) und Tailorshop (Putz-Osterloh 1981), und versucht deren psychometrische Schwächen durch die Nutzung von Itemsequenzen, also aufeinander basierende geschlossene Items, zu überwinden. Um den damit einhergehenden lokalen stochastischen Abhängigkeiten zu begegnen, wurde mit einer konstruktbasierten Testwertinterpretation gearbeitet, ohne auf sensiblere psychometrische Modelle zurückzugreifen.

Es wurden zwei Testprobleme entwickelt und, nach Pretest und Pilotierung, unter Studierenden von berufs- und wirtschaftspädagogischen, erziehungswissenschaftlichen, psychologischen und wirtschaftswissenschaftlichen Studiengängen erhoben. Nach einem Cleaning der Stichprobe konnten 207 Proband:innen für die Auswertung berücksichtigt werden. Ergänzend wurden die Testprobleme einer Untersuchung nach der Lautes-Denken-Methode unterzogen, wodurch unter Berücksichtigung von neun Einzelerhebungen ihre Validität nachgewiesen werden konnte.

Insgesamt lässt die Verteilung der ermittelten Testwerte auf ein angemessenes Niveau der Testprobleme schließen, da diesbezüglich Normalverteilungen festgestellt wurden. Für die Stichprobe bedeutet dies, dass Potenzial zur Entwicklung komplexen Problemlösens vorhanden ist. Aus begleitend erhobenen Skalen zu soziodemografischem Hintergrund, Einstellung zu Nachhaltigkeit, Leseinteresse sowie emotionaler Befindlichkeit konnten über Regressionsanalysen nur schwache Prädiktoren abgeleitet werden.

Insgesamt wurde ein Instrument entwickelt und zur Anwendung gebracht, das zwar mit einem hohen Entwicklungsaufwand einhergeht, jedoch in Erhebung und Auswertung effizienter ist als vergleichbare domänenbezogene Testprobleme mit offenen Elementen (bspw. bei Sembil 1982 und Seifried et al. 2016). Künftig wäre über eine Ausweitung der Erhebung, die Entwicklung weiterer Testprobleme mit anderem Domänenbezug oder die Adaption des Prinzips für teamorientiertes Problemlösen nachzudenken.

Entwicklung und Evaluation eines Kompetenztests auf Hochschulebene am Beispiel der organischen Chemie

Martin Steinbach, Carolin Eitemüller, Marc Rodemer, Maik Walpuski
Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund und Fragestellung

Im sekundären Bildungsbereich kommen der Modellierung und Messung von Kompetenzen eine zentrale Rolle zu. Kompetenzmodellierungen werden beispielsweise im Zuge internationaler Vergleichsstudien (z. B. PISA, TIMSS) eingesetzt um Bildungsstandards zu evaluieren und Bildungsprozesse sowohl auf System- als auch auf Unterrichtsebene zu verbessern (Bögeholz & Eggert, 2013). Analog zu den Entwicklungen im Schulkontext nimmt die Bedeutung der Kompetenzdiagnostik auch im Hochschulsektor immer weiter zu (Wolter & Schiener, 2016). Aufgrund der hohen Domänenspezifität des Kompetenzbegriffs (Klieme & Leutner, 2006) besteht hinsichtlich fachspezifischer Kompetenzmodelle und deren Übersetzung in kompetenzorientierte Testinstrumente noch erheblicher Entwicklungsbedarf. Dies gilt auch für das Studienfach Chemie und dessen Teildisziplin der organischen Chemie. Ausgehend von dem aufgezeigten Entwicklungsbedarf wurde im Zuge des Projekts ein Kompetenzstrukturmodell für die organische Chemie auf Hochschulebene entwickelt. Das Kompetenzmodell wurde anschließend durch die Entwicklung eines Testinstruments zur Erhebung des Fachwissens in der organischen Chemie operationalisiert. Zur Evaluation des Testinstruments werden folgende Forschungsfragen adressiert:

FF1: Bilden die Items des Fachwissenstests das entwickelte Kompetenzstrukturmodell zur organischen Chemie valide und reliabel ab?

FF2: Welchen Einfluss haben die Teildimensionen des für die organische Chemie entwickelten Kompetenzstrukturmodells auf die Aufgabenschwierigkeit von Leistungstestaufgaben?

Methoden

Bei der Entwicklung des Kompetenzstrukturmodells wurde auf eine dreidimensionale Struktur zurückgegriffen, da diese sich für das Abbilden naturwissenschaftlicher Kompetenzen bereits als zielführend erwiesen hat (Mayer & Wellnitz, 2013). Neben der Achse der kognitiven Prozesse und der Achse der Komplexität (vgl. Abb. 1) wird das für die organische Chemie spezifische Aufgabenmerkmal der Visualisierung mit in das Kompetenzstrukturmodell aufgenommen. Während bei der Achse der kognitiven Prozesse auf die Lernzieltaxonomie von Anderson und Krathwohl (2001) zurückgegriffen wird, orientiert sich die Achse der Komplexität an dem ESNaS-Kompetenzstrukturmodell (Walpuski et al., 2011), welches im Zuge der Evaluation der naturwissenschaftlichen Bildungsstandards verwendet wurde. In einem weiteren Schritt wurde das Kompetenzstrukturmodell durch die Entwicklung von N = 76 Items operationalisiert. Das Testinstrument wurde im Zuge einer Pilotierung erprobt und überarbeitet. In der anschließenden Hauptstudie wurde das Testinstrument an sieben deutschen Universitäten in Einführungsveranstaltungen der organischen Chemie eingesetzt. Insgesamt haben N = 357 Studierende den Fachwissenstest bearbeitet. Die Evaluation des Testinstruments fand auf Basis des Rasch- Modells statt. Die Passung zum Modell sowie die Güte des Tests wird durch die Betrachtung des Infits des wMNSQ- Werts sowie der Item- und Personen- Reliabilität sichergestellt. Um den Einfluss der Modellachsen auf die Aufgabenschwierigkeit zu bestimmen werden inferenzstatistische Tests durchgeführt.

Ergebnisse

Der wMNSQ-Wert des Testinstruments ($0.83 \leq wMNSQ \leq 1.24$; $-2.49 \leq t \leq 4.15$) weist eine gute Passung zum Rasch-Modell auf (Boone et al., 2014). Die Item-Reliabilität des Testinstruments liegt mit .98 in einem hervorragenden und die Personen-Reliabilität mit .80 in einem sehr guten Bereich.

Komplexität ($F(2,73) = 11.99$, $p < .001$, $\eta^2 = .25$) und kognitive Prozesse ($t(74) = -6.74$, $p < .001$, $d = -1.57$) konnten als schwierigkeitserzeugend identifiziert werden. Der Einfluss der Achse der Visualisierung wurde hinsichtlich der Aufgabenschwierigkeit nicht signifikant ($t(74) = 1.10$, $p = .275$). Bezogen auf die Aufgabenschwierigkeit ließ sich mit Hilfe des Kompetenzstrukturmodells eine Varianz von 43 % (korrigiertes $R^2 = .43$) aufklären.

Diskussion

Mit dem hier vorliegenden Testinstrument konnte eine Lücke bei der Entwicklung kompetenzorientierter Messinstrumente im Hochschulbereich für die organische Chemie geschlossen werden. Das Testinstrument kann zukünftig im Kontext der domänenspezifischen Diagnostik und somit für die Weiterentwicklung von chemiebezogenen Bildungsprozessen im Hochschulsektor eingesetzt werden. Das Projekt liefert zusätzlich methodische Einblicke in den Prozess der Kompetenztestentwicklung und dessen Evaluation auf Basis des Rasch- Modells.

E-05: Familiärer Hintergrund

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S06 S03 B06
Chair der Sitzung: Miriam Schwarzenthal

You can make it if you really try – or not? Parents' conceptions of social inequity and meritocracy and children's educational effort and aspirations

Miriam Schwarzenthal[1], Frederick de Moll[2]
[1]Bergische Universität Wuppertal; [2]Universität Bielefeld

Theoretical background

Most Western democratic societies cherish the promise of meritocratic status attainment. The idea of meritocracy is that social status is justly distributed based on prior accomplishments with schools as the engines of this process (Bills, 2019). However, children from families with low socioeconomic status and immigrant origin are profoundly disadvantaged at school (Clauss & Nauck, 2010; Skopek & Passaretta, 2021). Simultaneously, underprivileged groups are often optimistic about their educational opportunities (Lardier et al., 2019) and people in more unequal societies firmly believe that success is based on meritocratic competition and not on family-based advantages (Mijs, 2021). It remains unknown how such beliefs affect children's educational efforts and aspirations. Two mechanisms seem plausible: a) parents who believe in the meritocratic narrative that everybody can make it may instill a willingness to succeed in their children (Mijs et al., 2022). b) parents from disadvantaged groups who critically reflect on the lacking meritocracy and on social inequity might encourage their children to overcome obstacles (Heberle et al., 2020).

Research question

In this study, we address three main questions: (1) To what extent is a family's disadvantaged position (i.e., based on their socioeconomic status as well as their immigrant origin) associated with parents' beliefs about meritocracy and their critical reflection about social inequity in society? (2) How are a family's disadvantaged position, as well as parental beliefs about meritocracy and their critical reflection on social inequity related to children's educational efforts and aspirations? (3) Is the association between a family's disadvantaged position with children's educational efforts and aspirations (partly) mediated by parents' beliefs about meritocracy and their critical reflection on social inequity?

Method

The analyses draw on primary data of parents (n = 502) and children (n = 985, age: *M* = 8.9, SD = .79) at 16 primary schools in two major urban areas in Germany as part of the EDUCARE study (2010–2016). We conducted a path analysis with familial socioeconomic status (i.e., familial income, parental education) and immigrant origin as a predictor, parents' beliefs about meritocracy and their critical reflection of social inequity as a mediating variable, and children's educational effort and aspirations as a dependent variable. We also controlled for children's prior performance.

Results and their implications

There are three main findings: First, familial immigrant origin positively predicts parents' belief in meritocracy and children's educational aspirations and effort. The latter is in accordance with research showing that children of immigrants display remarkable academic effort (Friberg, 2019). However, the effect of immigrant origin on children's outcomes is not mediated by meritocratic beliefs. Second,

parental education correlates positively with parents' critical reflection of social inequity. Third, parents' critical reflection of social inequity is unrelated to children's educational effort and aspirations, whereas parents' meritocratic beliefs predict children's aspirations above and beyond their socioeconomic status and immigrant origin.

Our study merges literature on the critical reflection of social inequities (e.g., Heberle et al., 2020) with recent scholarship on meritocracy beliefs (Mijs et al., 2022). Our findings suggest that awareness of unequal opportunities in the educational system is more common among more privileged parents. This may be explained with differential opportunities to learn about and reflect on structural inequities (e.g., Flanagan et al., 2014; Kornbluh et al., 2019). Overall, parents who attribute success to individual effort encourage children's educational effort and aspirations irrespective of their social or immigrant origin. However, it remains an open question how parents' beliefs about meritocracy and their critical reflection on social inequity are not just related to children's academic effort and aspirations, but also to actions that explicitly aim to reduce social inequities (e.g., Diemer et al., 2021).

Stereotypenbedrohung für Schüler:innen mit türkischem und arabischem Migrationshintergrund in einer Wortlernaufgabe. Stellt das Fähigkeitsselbstkonzept im Fach Deutsch eine Ressource dar?

Johanna Hildebrandt[1], Madita Frühauf[1], Sabrina König[2], Justine Stang-Rabrig[2], Bettina Hannover[1]
[1]FU Berlin; [2]TU Dortmund

Lern- und Leistungserfolge von Mitgliedern ethnischer Minderheiten sinken, wenn ein negatives fähigkeitsbezogenes Stereotyp über ihre Gruppe ausgesprochen wird (vgl. Appel et al., 2015), ein als Stereotypenbedrohung bezeichneter Effekt (Steele & Aronson, 1995). Dieser konnte für türkischstämmige Schüler:innen in Deutschland bereits nachgewiesen werden (Froehlich et al., 2016; Martiny et al., 2014). Eine weitere möglicherweise von Stereotypenbedrohung betroffene Gruppe in Deutschland stellen Menschen arabischer Herkunft dar: Froehlich und Schulte (2019) fanden, dass das Stereotyp über Zuwander:innen aus arabischsprachigen Ländern durch noch geringere Wärme- und Kompetenzzuschreibungen charakterisiert ist als das über türkischstämmige Menschen. Wir haben den Effekt der Stereotypenbedrohung für türkischstämmige und arabischstämmige Schüler:innen untersucht, um somit zu einer Erklärung ihrer Benachteiligung im deutschen Bildungssystem (vgl. Weis et al., 2019) beizutragen.

Studien zeigen, dass Schüler:innen mit Migrationshintergrund nach Berücksichtigung ihrer Kompetenzen anspruchsvollere Bildungsgänge wählen (Überblick: Dollmann, 2021; Hadjar & Scharf, 2019) und an diesen häufiger zu finden sind als Schüler:innen ohne Migrationshintergrund (für Deutschland Kristen & Dollmann, 2010; Lenz et al., 2021; Relikowski et al., 2012). Eine mögliche Erklärung ist eine besonders günstige motivationale Konstellation, die durch starke Aspirationen und Selbstkonzepte charakterisiert und mit adaptivem Lernverhalten (Marsh & O'Mara, 2008; Trautwein & Möller, 2016) sowie positiven Lern- und Leistungsergebnissen verbunden ist (Marsh & Craven, 2006). So fanden beispielsweise Hildebrandt (2014) sowie Tiedemann und Billmann-Mahecha (2006), dass türkische Schüler:innen bei vergleichbaren Leistungen höhere Fähigkeitsselbstkonzepte aufwiesen als Schüler:innen ohne Migrationshintergrund. Wir untersuchten das Fähigkeitsselbstkonzept im Fach Deutsch als eine potentielle Ressource, die türkischstämmigen und arabischstämmigen Schüler:innen

helfen kann, eine Stereotypenbedrohung zu bewältigen.
In einer experimentellen Studie haben wir 127 Schüler:innen mit türkischer und 53 Schüler:innen mit arabischer Familiensprache aus neunten und zehnten Klassen untersucht. In einem Prätest-Posttest-Kontrollgruppendesign bearbeiteten diese individuell im Klassenkontext auf einem Tablet eine Aufgabe, bei der sie schwierige deutsche Wörter lernen mussten. Zuvor sahen sie randomisiert eines von vier Videos, in dem eine Lehrerin auf Unterschiede im Erfolg in der Wortlernaufgabe zwischen Schüler:innen mit türkischer (für türkischstämmige Versuchspersonen) bzw. nicht-deutscher Familiensprache (für arabischstämmige Versuchspersonen) hinwies (Stereotypenaktivierung) oder aber sagte, dass es keine solchen Unterschiede gäbe (keine Stereotypenaktivierung). Gekreuzt mit dieser Manipulation wurde der Name der Lehrkraft (türkisch versus deutsch) variiert.
Wir erwarteten einen negativen Effekt der Stereotypenbedrohung auf den Wortschatzerwerb bei türkischstämmigen und arabischstämmigen Schüler:innen und dass ein hohes Fähigkeitsselbstkonzept im Fach Deutsch einer Stereotypenbedrohung entgegenwirken kann. Weiter haben wir exploriert, inwiefern die Herkunft der Lehrkraft den Lernzuwachs unter den beiden Bedingungen der Stereotypenbedrohungsbedingungen differentiell beeinflusst. In multiplen Regressionsanalysen untersuchten wir die Posttestwerte in den vier Versuchsbedingungen und kontrollierten für die Prätestwerte, die Anzahl von Büchern zuhause, das Geschlecht, die Deutschnote, die Schulform und das Fähigkeitsselbstkonzept unter Verwendung der Analysesoftware Mplus Version 8.7 (Muthén & Muthén, 1998-2017).
Deskriptiv zeigten sich für arabischstämmige Schüler:innen vergleichsweise geringere Prä- und Posttestwerte als für türkischstämmige. Inferenzstatistisch zeigten sich keine Effekte der experimentellen Variationen " Lehrkraftherkunft" und "Stereotypenbedrohung" auf den Lernzuwachs. Für türkischstämmige Schüler:innen ergab sich ein positiver Effekt des Fähigkeitsselbstkonzepts auf den Wortlernzuwachs (.31***) und für den Interaktionsterm Fähigkeitsselbstkonzept x Stereotypenbedrohung, dass sie bei hohem Selbstkonzept bei Stereotypenbedrohung weniger hinzulernten (-.24**). Für arabischstämmige Schüler:innen zeigte sich ein umgekehrtes Befundmuster mit negativen Effekten ausgehend vom Fähigkeitsselbstkonzept (-.21+); dafür lernten sie bei hohem Selbstkonzept und Stereotypenbedrohung mehr hinzu (.41***).
Eine vorsichtige Interpretation dieser unerwarteten Effekte ist, dass türkischstämmige Schüler:innen mit hohem Selbstkonzept mit der Domäne Deutsch stark identifiziert sind, wobei Domänenidentifikation Stereotypenbedrohungseffekte hervorrufen kann (z.B. Keller, 2007; Steele, 1997). Im Unterschied zu den türkischstämmigen Jugendlichen wurde die arabischstämmige Gruppe nicht direkt angesprochen ("Jugendliche mit einer nichtdeutschen Familiensprache"). Dies hat möglicherweise dazu geführt, dass sich die Schüler:innen mit hohem Fähigkeitsselbstkonzept bei dem Hinweis auf Unterschiede zur deutschsprachigen Gruppe besonders motiviert fühlten, beim Lernen neuer deutscher Wörter gut abzuschneiden.

Selbstberichtete Gründe für die (Nicht-)Nutzung von Familiensprachen beim mehrsprachigkeitssensiblen Reziproken Lehren

Valentina Reitenbach[1], Jasmin Decristan[1], Dominique Rauch[2,4], Victoria Bertram[2,3], Katharina Schneider[2]
[1]Bergische Universität Wuppertal; [2]Pädagogische Hochschule Ludwigsburg; [3]Goethe-Universität Frankfurt am Main; [4]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Der Einbezug der Familiensprachen von Schüler*innen im Unterricht ist nicht nur politisch (Council of Europe, 2001) gefordert, sondern lässt sich auch anerkennungstheoretisch (Honneth, 1992) begründen und kann Sprachbarrieren bei der Kommunikation überbrücken (z.B. durch Code-Switching beim Nutzen verschiedener Sprachen; Moschkowitsch, 2007). Gerade kooperative Lehr-Lernsettings wie das Reziproke Lehren (RL; Rosenshine & Meister, 1994) können einen Rahmen für mehrsprachige Interaktion bieten, ohne dass die Lehrkräfte die Sprachen der Schüler*innen selbst sprechen müssen. Dennoch scheinen Schüler*innen ihre Familiensprachen hierbei kaum zu nutzen (z.B. Meyer & Prediger, 2011).

Aus Grosjeans (2020) Theorie zu den Sprachmodi bilingualer Sprecher*innen lassen sich hierfür drei übergeordnete Gründe ableiten: das Individuum selbst, die Sprachpartner*innen und der Kontext. Eine erste empirische Bestätigung hierfür findet in den Analysen einer kleinen Stichprobe mit deutsch-türkischsprachigen Kindern im Rahmen einer außerunterrichtlichen Peer-Tutoring-Intervention, in der die Schüler*innen verschiedene Impulse zur mehrsprachigen Interaktion erhielten (Schastak et al., 2017). Aktuell findet sich keine Studie im Unterrichtssetting und mit einer umfangreichen Stichprobe.

Der vorliegende Beitrag adressiert diese Forschungslücke und untersucht anhand einer umfangreichen Stichprobe in Grundschulklassen, welche Gründe die 435 mehrsprachigen Viertklässler*innen mit ganz unterschiedlichen Familiensprachen für die Nutzung und die Nicht-Nutzung ihrer Familiensprachen während einer RL-Intervention angaben. Das RL umfasste dabei spezifische Impulse zur Nutzung mehrsprachiger Interaktion.

Vor der Intervention nahmen die Lehrkräfte an drei Nachmittagen an einer Fortbildung zu den Themen RL und Mehrsprachigkeit im Unterricht teil. Die Inhalte waren in eine Unterrichtsreihe zur Leseförderung im Deutschunterricht eingebettet, die die Lehrkräfte anschließend in ihren Klassen (N=37) umsetzten. Nach der Instruktion durch die Lehrkraft, übten die Schüler*innen in Kleingruppen Lesestrategien nach der Methode des RL ein. Die Kleingruppen wurden so zusammengesetzt, dass möglichst Kinder mit den gleichen Familiensprachen in einer Gruppe waren. Mehrsprachige Interaktion wurde mit drei Impulsen eingeführt: 1. Schaffung eines mehrsprachenfreundlichen Klassenklimas (Sprachenporträts; Gogolin & Neumann, 1991); 2. mehrsprachiges Unterrichtsmaterial; 3. Gruppenbildung auf Grundlage gemeinsamer Familiensprachen.

Nach der Unterrichtseinheit wurden die Schüler*innen in einem Fragebogen gefragt, ob sie während der Unterrichtseinheit eine andere Sprache als Deutsch gesprochen haben (ja/nein) und warum (offene Frage). Die Angaben wurden mithilfe der deduktiven qualitativen Inhaltsanalyse nach Mayring (2000) kodiert. Den Ausgangspunkt bildete das Kategoriensystem von Schastak et al. (2017). Etwa 33% des Materials wurde doppelt kodiert, die Interrater-Reliabilität war hoch (κ_n = .75-.95; Brennan & Prediger, 1981). Sowohl für die Nutzung als auch für die Nicht-Nutzung von mehrsprachiger Interaktion wurden vier Hauptkategorien identifiziert: *Individuum* ("Weil es mir gefallen hat"), *Gruppe* ("Weil es uns gefallen hat"), *Material* ("Wegen der Karten") und *Sprachangebot* ("Weil wir es durften"). *Zeit* war ein zusätzlicher Grund für die Nicht-Nutzung ("Weil wir keine Zeit hatten"). Innerhalb der individuellen und gruppenbezogenen Gründe wurden zwei Unterkategorien identifiziert *Kompetenzen*

("Weil ich nicht so gut Tamil spreche"); *Emotionen* ("Weil es Spaß macht"). Die Kategorie *Sprachangebot* beinhaltete *Angebot*, *Aufforderung* oder *Verbot* als Unterkategorien.

Von 435 mehrsprachigen Schüler*innen haben 374 bis zu drei Gründe für die (Nicht-)Nutzung ihrer Familiensprache während der Intervention angegeben. Insgesamt wurden 432 Gründe kodiert. Sowohl für die Nutzung (n=283) als auch für die Nicht-Nutzung (n=149) ihrer Familiensprache nannten Schüler*innen hauptsächlich individuelle (Nutzung: n=133; Nicht-Nutzung: n=57) und gruppenbezogene (Nutzung: n=96; Nicht-Nutzung: n=69) Gründe. Schüler*innen, die eine Nutzung angaben, taten dies mehrheitlich aus Freude (n=64) oder weil sie sich oder die Gruppe als in der Sprache kompetent wahrnahmen (n=155). Zentral für eine Nicht-Nutzung waren der eigene Mangel an Sprachkompetenz sowie derjenige der Gruppenmitglieder (n=93).

Die Vielfalt der angegebenen Gründe verdeutlicht die Individualität der Entscheidungen der Schüler*innen für oder gegen mehrsprachige Interaktion. Dies zeigt zum einen, dass das Angebot mehrsprachiger Interaktion nicht für alle Schüler*innen geeignet ist, und zum anderen, dass es für die Stimulierung mehrsprachiger Interaktion gilt, auf unterschiedliche Motive und Barrieren einzugehen.

Epistemologische Überzeugungen von Migrant/-innen: Zur Bedeutung von Partizipation an Kulturräumen

Caroline Rau
Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund:

Auf der einen Seite kann unter wissenschafts- bzw. erkenntnistheoretischen Perspektiven danach gefragt werden, unter welchen Bedingungen Wissen bzw. unterschiedliche Wissensformen generiert werden. Auf der anderen Seite kann auf Basis eines empirischen Zugangs gefragt werden, inwiefern bzw. wie Akteur/-innen Wissen hervorbringen und welche unterschiedlichen Wissensformen ihr Handeln orientieren. Diese beiden Facetten werden im Kontext des Forschungsprogramms „Epistemologische Überzeugungen“ aufgegriffen. Sie umfassen „jene Vorstellungen und subjektive Theorien […], die Personen über das Wissen und den Wissenserwerb generell oder in spezifischen Domänen entwickeln“ (Baumert und Kunter 2006, S. 498). Hierunter fallen „subjektive Überzeugungen über die Struktur, Verlässlichkeit, Genese, Validierung und Rechtfertigung von Wissensbeständen“ (ebd.). Epistemologische Überzeugungen wurden bislang vornehmlich aus einer domänenspezifischen bzw. domänenunabhängigen Perspektive (z.B. mathematisch-naturwissenschaftlicher oder geisteswissenschaftlicher Fächerkanon, Pädagogik, Psychologie, Fachdidaktik usw.) und in der Perspektive von Lebensalter sowie Expert/-innen bzw. Noviz/-innen (z.B. Lehramtsstudierende, Referendar/-innen, Lehrkräfte) erforscht (vgl. für einen ausführlichen Forschungsüberblick Rau 2020a, 2020b, 2021). Ein Forschungsdesiderat stellt die Kontext(un)abhängigkeit epistemologischer Überzeugungen dar (vgl. Mayer und Rosman 2016): also beispielsweise die kulturelle Affinität von Personen in Bezug auf Wissensproduktion und -erwerb. Gleichzeitig wurde lange Zeit die empirisch nicht mehr haltbare Annahme vertreten, wonach „die Produktion und Anwendung von Wissen[…] ein Prozess sei, der nichts mit räumlichen Gegebenheiten zu tun habe. Wissen sei [schließlich] universell, es könne überall generiert werden und sei ubiquitär verfügbar“ (Meusburger 2006, S. 269). Damit bescheinigt Meusburger dem Forschungsprogramm

„raumblinde[…]' Ansätze" (ebd.). Führt man beide Forschungsdesiderate - kulturbezogene Kontextabhängigkeit und Raumbezug - zusammen, so ist anzunehmen, dass der „Kulturraum als Areal [zu sehen] ist, in dem bestimmte Leitbilder oder ‚Wahrheiten' dominieren" (vgl. ebd, S. 283-285). Interessanterweise wird die relationale Verbindung von Raum und Wissensgenese zunehmend im Themenspektrum der Science of Science in den Blick genommen (vgl. Brunner 2020; Wöpking 2016; Müller und Scholz 2012; Latour 1994; Knorr-Cetina 1984; Knorr-Cetina 2002). Die Anlehnung an die vorausgegangenen Überlegungen lässt es den Schluss zu, dass eine empirische Untersuchung der kulturraumbezogenen Kontextabhängigkeit von epistemologischen Überzeugungen vielversprechende Befunde hervorbringen dürfte.

Deshalb greift der vorliegende Beitrag die *Forschungsdesiderate* auf. Der Fokus liegt dabei auf Personen, die aufgrund von Globalitätserfahrungen an unterschiedlichen Kulturräumen partizipieren: *Wie gestalten Migrant/-innen die Genese und den Erwerb von Wissen vor dem Hintergrund ihrer kulturräumlichen Erfahrungswelt? Über welche epistemologischen Überzeugungen verfügen Migrant/-innen, die an unterschiedlichen Kulturräumen aufgrund ihrer Globalitätserfahrungen partizipieren?* Die Forschungsfrage schließt auch an bisherige Untersuchungen an, wonach u.a. Globalitätserfahrungen, die in unterschiedlichen kulturellen Räumen gesammelt werden, Einfluss auf weltgesellschaftliches Lernen haben (vgl. Scheunpflug 2003, 2011).

Methode: Die Fragestellung dieser Untersuchung wurde mit einem qualitativ-rekonstruktiven Forschungsdesign – im Besonderen mit der Dokumentarischen Methode (vgl. Bohnsack 2014, 2017) – im Rahmen einer explorativen Sekundäranalyse bearbeitet. Als Datengrundlage dienten narrative Interviews (n=22), die mit Migrant/-innen geführt wurden. Rekurrierend auf das „Theoretical Sampling"-Verfahren (vgl. Glaser und Strauss 2010) wurde das gesättigte Sample mittels unterschiedlicher Samplingstrategien zusammengestellt. Diese sind: 1., 2., 3. Migrationsgeneration mit Bezug zur Türkei; Bildungshintergrund; Religionszugehörigkeit. Im Kontext der Datengeneralisierung wurden die Befunde zu einer Typologie verdichtet (vgl. Nentwig-Gesemann 2013; Schäffer 2022).

Ergebnisse und ihre Bedeutung: Bei der Auswertung des Materials zeigte sich, dass der Kulturraumbezug im Kontext von Wissensgenese und Wissenserwerb für Migrant/-innen eine besondere Rolle spielt, wovon deren epistemologische Überzeugungen berührt sind. Im Besonderen wird um die Suche nach epistemischer Autorität gerungen. Gleichzeitig machen die Befunde deutlich, dass Personen, die an unterschiedlichen Kulturräumen partizipieren, unterschiedliche Modi im Umgang mit epistemischer Autorität inkorporiert haben. In unterschiedlichen Kulturräumen wird zwischen unterschiedlichen Modi in Bezug auf epistemische Autorität changiert. Die Ergebnisse bieten wertvolle Anregungen für die Lehr-/Lernforschung: Sie zeigen, dass bei der Sensibilisierung von Lernenden für die epistemischen Grundlagen der Wissensproduktion und des Wissenserwerbs die kulturräumliche Gebundenheit der Lernenden berücksichtigt werden sollte.

E-06: Lernen, Kognition und Enthusiasmus

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S06 S04 B06
Chair der Sitzung: Marco Rüth

Online-Lernen während der Covid-19-Pandemie: Wie Wahrnehmung, Engagement und Leistung von Studierenden an Hochschulen mit ihren persönlichen Eigenschaften zusammenhängen

Marco Rüth, Kai Kaspar
Department Psychologie, Universität zu Köln

Theoretischer Hintergrund

Studierende an Hochschulen standen während der Covid-19-Pandemie vor unerwarteten Herausforderungen beim Online-Lernen (Hoss et al., 2022; Radu et al., 2020). Die Vorteile des Online-Lernens sind bekannt, doch die Online-Lernerfahrungen können zwischen Studierenden unterschiedlich ausfallen. Auch kann es stark von ihren persönlichen Eigenschaften abhängen, wie gut sich Studierende an solche Lebensveränderungen anpassen (Caspi & Moffitt, 1993; Pinquart & Silbereisen, 2004). Daher muss die Rolle, die persönliche Merkmale der Studierenden bei ihren Online-Lernerfahrungen spielen, noch genauer untersucht werden.

Fragestellung

Diese korrelative Querschnittsstudie untersucht, wie persönliche Eigenschaften von Studierenden an Hochschulen mit ihren Online-Lernerfahrungen während der Covid-19-Pandemie zusammenhängen. Basierend auf theoretischen Ansätzen wie der sozial-kognitiven Lerntheorie (Bandura, 2006) und empirischen Befunden vermuteten wir Zusammenhänge zwischen Online-Lernerfahrungen (fünf Dimensionen von Wahrnehmung des Online-Lernens, Engagement in Online-Kursen und Leistung in Online-Kursen) und den folgenden persönlichen Eigenschaften: Alter, Geschlecht und Studiensemester (H1a-c), Big Five-Persönlichkeitsmerkmale (H2a-e), Selbstregulationsfähigkeiten (H3), allgemeine und akademische Selbstwirksamkeit sowie Selbstwirksamkeit in Bezug auf den Umgang mit digitalen Medien (H4a-c) und Zustandsangst und Angst in Bezug auf die Covid-19 Pandemie (H5a und H5b).

Methode

Anhand einer Online-Befragung erhoben wir vollständige Angaben zu persönlichen Merkmalen und Erfahrungen mit dem Online-Lernen von 413 Studierenden an deutschen Hochschulen im Alter zwischen 18 und 61 Jahren (M_{Alter} = 25.47, SD_{Alter} = 7.18, davon 354 weiblich). Die meisten Teilnehmenden waren Studierende der Psychologie und gesundheitsbezogener Fächer. In der Online-Befragung kamen verschiedene Skalen mit guten psychometrischen Eigenschaften zum Einsatz. Mit Hilfe multipler Regressionsanalysen untersuchten wir, wie Alter, Geschlecht und Studiensemester der Studierenden, die Big Five-Persönlichkeitsmerkmale, die Selbstregulationsfähigkeiten, die verschiedenen Facetten der Selbstwirksamkeit und die Angst mit ihren Online-Lernerfahrungen zusammenhängen.

Ergebnisse und ihre Bedeutung

Insgesamt ergab sich durch die Berücksichtigung aller persönlichen Eigenschaften (H1-H5) eine Varianzaufklärung zwischen 10% und 22% für die fünf Dimensionen von Wahrnehmung des Online-Lernens, von 22% für das selbsteingeschätzte Engagement in Online-Kursen und von 29% für die

selbsteingeschätzte Leistung in Online-Kursen. Es zeigte sich, dass das Alter der Studierenden in einem signifikant positiven Zusammenhang mit dem Engagement in Online-Kursen und allen fünf Dimensionen der Wahrnehmung des Online-Lernens stand (H1a). So stimmten ältere Studierende eher zu, dass Online-Lernen Zugang zu kostenlosen und unbegrenzten Lernmaterialien und Ressourcen bietet, Interaktionen mit Dozierenden und unter Studierenden ermöglicht, die Lernzeiten, -orte und -prozesse selbst bestimmt werden können, Wissenserwerb in den gewünschten Kompetenzbereichen fördert und weniger Stress und Belastungen bedeuten. Unsere Ergebnisse bestätigen auch, dass Selbstregulationsfähigkeiten (H3), akademische Selbstwirksamkeit (H4b) und Selbstwirksamkeit in Bezug auf den Umgang mit digitalen Medien (H4c) wichtige Faktoren bei verschiedenen Online-Lernerfahrungen sind. Im Gegensatz dazu waren die Big Five-Persönlichkeitsmerkmale der Studierenden (H2) und ihre Angst (H5) für die meisten Online-Lernerfahrungen weniger wichtig. Bemerkenswert ist, dass sich mehrere bivariate Zusammenhänge zwischen persönlichen Eigenschaften und Online-Lernerfahrungen nicht in dem multiplen Regressionsmodell widerspiegeln, welches alle Variablen gleichzeitig berücksichtigt. Dies unterstreicht, dass einzelne Variablen und Zusammenhänge nicht nur isoliert betrachtet werden sollten.

Aus praktischer Sicht ergeben sich durch die Identifikation von zentralen Persönlichkeitseigenschaften wertvolle Ansatzpunkte für zukünftige Bildungsangebote. So könnten beispielsweise Selbstregulationsfähigkeiten, akademische Selbstwirksamkeit und Selbstwirksamkeit in Bezug auf den Umgang mit digitalen Medien in Online-Lernsettings durch Videovorlesungen, Quizspiele oder andere digitale Tools zum Selbststudium und kollaborativen Lernen gefördert werden (z. B. Jansen et al., 2020; Pérez-Álvarez et al., 2018; Rüth et al., 2021). Gleichzeitig scheint es weniger effektiv zu sein, Ressourcen in die Berücksichtigung anderer persönlicher Eigenschaften zu investieren, wie etwa eine je nach Geschlecht der Studierenden unterschiedliche Ausgestaltung von Online-Lernumgebungen (Yu, 2021). Insgesamt erweitern diese Ergebnisse bekannte Vorteile von Online-Lernen auf Studienmotivation und -erfolg um Erkenntnisse zur wahrgenommenen Qualität des Online-Lernens und der Rolle von persönlichen Eigenschaften während der Covid-19-Pandemie.

Kultuturdidaktische Kognitionen von Englischlehrpersonen zu interkulturellen Unterrichtsmerkmalen und Critical Incidents aus Englischlehrwerken – Eine strukturierende qualitative Inhaltsanalyse

Bernd Frankemölle, Kerstin Göbel
Universität Duisburg-Essen, Deutschland

Englischlehrpersonen erachten die interkulturelle Kompetenzvermittlung mehrheitlich als sinnvoll, gleichzeitig priorisieren sie häufig die Zielsprachenförderung (z.B. Oranje & Smith, 2018; Sercu, 2005b; 2006). Auch die Vieldimensionalität von interkultureller Kompetenz scheint Fremdsprachenlehrpersonen bewusst. Wird interkulturelles Lernens jedoch im Unterricht adressiert, stehen häufig Fakteninformationen bzw. die Wissensdimension im Vordergrund (z.B.; Brunsmeier, 2017; Göbel & Helmke, 2010; Kreft, 2020; LACE, 2007; Sercu, 2005b; 2006). Lehrpersonenzentrierung und wenig Schüleraktivierung sind bisweilen die Folge (z.B. Göbel, Lewandowska & Diehr, 2017; Gonen & Saglam, 2012; LACE, 2007). Die aktuelle Befundlage weist

auf vielschichtige Postionierungen von Englischlehrpersonen gegenüber der Vermittlung interkultureller Kompetenz im Englischunterricht hin (Petravic, 2015). Aus Perspektive der Aufgabenorientierung gibt es bislang jedoch kaum Befunde zu den professionellen Einschätzungen von Englischlehrpersonen hinsichtlich interkultureller Unterrichtsaufgaben (Van den Branden, 2016).

Welche professionellen Perspektiven Englischlehrpersonen (Blömeke et al., 2015; Legutke & Schart, 2016) auf interkulturelle Aufgabenmaterialien haben und inwiefern ein Zusammenhang zwischen allgemeinen und situationsspezifischen kulturdidaktischen Einschätzungen besteht, stellt entsprechend das Erkenntnisinteresse dieser Studie dar. Hierfür wurde unter Berücksichtigung bildungspolitischer Anforderungen an die interkulturelle und fachbezogene Lehrpersonenbildung (KMK, 2019a; 2019b) sowie professions-, handlungs- und kompetenztheoretische Perspektiven (Blömeke et al., 2015; Esser, 1996; Legutke & Schart, 2016; Petravic, 2015; Sercu, 2005a) ein *Modell zur Analyse kulturdidaktischer Kognitionen von Englischlehrpersonen zum interkulturellen Fremdsprachenunterricht und zu interkulturellen Critical Incident-Aufgaben* entworfen. Es dient als theoretische Orientierung für eine strukturierende qualitative Inhaltsanalyse (Kuckartz, 2018) von Leitfadeninterviews mit 17 gymnasialen Englischlehrpersonen. Sie haben zwischen 3 und 30 Jahren Berufserfahrung und stehen unterschiedlich häufig in Kulturkontakt (von „jede Woche“ bis „gar nicht (mehr)“).

Der Interviewleitfaden wurde orientiert an Merkmalen des problemzentrierten und des fokussierten Interviews (Merton, Fiske & Kendal, 1990; Witzel, 1985). Es wurde zunächst über allgemein relevante, dann über spezifisch interkulturell bedeutsame Unterrichtsaspekte und -erfahrungen und schließlich über Einschätzungen zum Potential von interkulturellen Critical Incidents (CIs) für interkulturelles Lernen gesprochen. Hierfür wurden zwei CI-Aufgaben als Stimuli eingesetzt, die sich in einer Lehrwerkanalyse von CIs in 30 gymnasialen Englischlehrwerken hinsichtlich adressierter interkultureller Kompetenzdimensionen, Aufgabenprinzipien sowie kultureller Merkmale als kontrastreich herausgestellt haben.

In der Interviewauswertung wurden *allgemeine* Vorstellungen zu interkultureller Kompetenz sowie zu Merkmalen guten interkulturellen Englischunterrichts von konkret erlebten also *handlungsnahen* interkulturellen Unterrichtserfahrungen und situativen Einschätzungen zum Potential der CIs für interkulturellen Englischunterricht unterschieden (Leuchter et al., 2006; OECD, 2009). Weiterhin wurden interkulturell relevante Haltungen und Motivationen (Petravic, 2015; Sercu, 2005a) in den Interviews codiert, um affektive Färbungen der geäußerten Vorstellungen herauszustellen. Die Kategorien sind thematisch aufgebaut und wurden deduktiv-induktiv entwickelt (Kuckartz, 2018).

Das Datenmaterial wurde in einem ersten Durchgang codiert und wird derzeit zweitcodiert.

Erste Analysen zeigen, dass 75% der befragten Lehrpersonen komplexe Vorstellungen von interkultureller Kompetenz äußern, welche neben der Wissensdimension auch affektive und / oder verhaltensbezogene Facetten einschließen. In berichteten Unterrichtserfahrungen mit interkulturellem Englischunterricht sowie in Äußerungen zu den CI-Aufgaben wird jedoch in 50% der Fälle die Wissensdimension betont, wohingegen die Verhaltensdimension nur in 37,5% und die affektive Dimension sogar nur in 12,5% der Fälle adressiert wird. In ähnlicher Weise äußern die Lehrpersonen

differenziertere allgemeine Vorstellungen von gutem interkulturellem Unterricht (Nennung von durchschnittlich 5-6 relevanten didaktischen Aspekten). 87,5% der Lehrpersonen beschreiben die Bedeutung von Interaktions- und Erfahrungsorientierung. In den Berichten eigener Unterrichtserfahrungen und den Äußerungen zu den CIs werden deutlich weniger Aspekte als relevant angesehen. Erfahrungs- und Interaktionsorientierung werden hier seltener adressiert (37,5% der Fälle). Die Daten zeigen, dass kulturdidaktische Vorstellungen je nach Bezugskontext variieren können und bekräftigen, dass interkulturelle Perspektiven für das eigene Unterrichtshandeln noch deutlicher durch die Adressierung in Lehrwerken sowie in der Lehrkräfteaus- und -fortbildung unterstützt werden müssen (Göbel & Helmke, 2010).

Zum Zeitpunkt des Vortrags werden die Interviews final codiert und das Kategoriensystem abgesichert sein. Der Beitrag dient der Präsentation und Diskussion ausgewählter Befunde dieser Interviewstudie.

Enthusiasmus von Mathematikstudierenden – Ein Konstrukt mit mehreren Facetten

Anika Radkowitsch, Daniel Sommerhoff
IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften und der Mathematik, Germany

Enthusiasmus von Lehrkräften ist eine wichtige Komponente von Unterrichtsqualität. Enthusiasmus wird zumeist als affektive Variable beschrieben, die das innere Erleben von Freude und Spaß der Lehrkraft als eine relativ überdauernde Personeneigenschaft (Trait) beschreibt und sich in spezifischem Verhalten (State) manifestiert (Keller et al., 2016; Kunter et al., 2011). Da sich affektiv-motivationale Variablen auf einen spezifischen Gegenstand beziehen (z.B. Deci & Ryan, 2000), ist Enthusiasmus objektspezifisch zu operationalisieren. Entsprechend wurde in der Vergangenheit zwischen Enthusiasmus einer Lehrkraft für das zu unterrichtende Fach und für das Unterrichten des Fachs als Aktivität unterschieden (Kunter et al., 2011). Es konnte gezeigt werden, dass sich beide Facetten empirisch trennen lassen und dass der Enthusiasmus für das Fach im Gegensatz zum Enthusiasmus für das Unterrichten die Unterrichtsqualität kaum beeinflusst (Kunter et al., 2008). Dies ist teilweise überraschend, da Theorien zur intrinsischen Motivation implizieren, dass der Enthusiasmus für ein Fach die Auseinandersetzung mit diesen Inhalten verstärken sollte (Deci & Ryan, 2000) und so beispielsweise die Unterrichtsvorbereitung positiv beeinflussen könnte. Ein Erklärungsansatz für die Befunde ist, dass bisher der Enthusiasmus für das Fach mit globalen Skalen (zumeist zum Fach Mathematik als Ganzes) gemessen wurde. Ufer et al. (2017) zeigten im Kontext von Interesse jedoch, dass sich der grundlegende fachliche Unterschied zwischen Schulmathematik und universitärer Mathematik auch im Interesse von Lehramtsstudierenden widerspiegelt, was nahelegt, dass sich auch der Enthusiasmus von Lehrkräften und Lehramtsstudierenden hinsichtlich beider Bereiche unterscheidet.

Die vorliegende Studie untersucht daher, inwiefern sich neben Enthusiasmus für das Unterrichten und das Fach auch spezifischere Facetten des Fachenthusiasmus für Mathematik unterscheiden lassen und inwiefern diese von der allgemeinen Fachenthusiasmus-Skala abgebildet werden. Dafür wurden 301 Mathematiklehramtsstudierende gebeten ihren Enthusiasmus für das Fach Mathematik im Allgemeinen (4 Items, *alpha* = .74), für Schulmathematik (5 Items, *alpha* = .91), für universitäre Mathematik (5 Items, *alpha* = .93) und für das Unterrichten von Mathematik (5 Items, *alpha* = .96) auf

einer 5-stufigen Likert-Skala einzuschätzen. Zur Beantwortung der Forschungsfragen wurden ein konfirmatorisches Messmodell mit einer vier-Faktor-Struktur sowie ein Strukturmodell mit Schulmathematik und universitärer Mathematik als Subfaktoren von Mathematik im Allgemeinen mit dem *lavaan* Paket in R erstellt und empirisch verglichen.

Die Analysen zeigten einen guten Fit für die 4-Faktor-Struktur ($Chi^2(171) = 295{,}40$, $p < .001$; CFI = .968; RMSEA = 0.058, $p = .076$) wobei Enthusiasmus für Mathematik im Allgemeinen stark mit Enthusiasmus für universitäre Mathematik ($r = .56$*) und gering mit Enthusiasmus für Schulmathematik ($r = .24$*) korrelierte. Enthusiasmus für universitäre Mathematik und für Schulmathematik korrelierten nicht ($r = -.02$, *ns*). Der Enthusiasmus für Unterrichten korrelierte nicht oder nur gering mit allen anderen Skalen ($r < .24$). Auch das Strukturmodell zeigte einen passablen Fit (Chi^2 (147) = 307,517, $p < .001$; CFI = .965; RMSEA = 0.060, $p = .038$). Das Strukturmodell fittete die Daten jedoch signifikant schlechter als das reine Messmodell ($Chi^2(1) = 12.114$, $p < .001$).

Die Ergebnisse bestätigen, dass sich der Enthusiasmus für die universitäre Mathematik und Schulmathematik unterscheiden (s. Ufer et al., 2017) und nicht signifikant miteinander korrelieren. Auch bestätigen die Ergebnisse, dass sich Enthusiasmus für Mathematik im Allgemeinen und für das Unterrichten trennen lassen. Weiter zeigt sich, dass sich der Enthusiasmus für Mathematik im Allgemeinen nicht als übergeordneter Faktor des Enthusiasmus in den spezifischeren Kontexten eignet. Diese Befunde legen nahe, dass der Enthusiasmus für das Fach Mathematik im Allgemeinen ein (zu) unspezifischer Indikator für den Enthusiasmus für das Fach Mathematik in einem spezifischen Kontext ist und damit könnten Befunde von Kunter et al. (2008) zum geringen Einfluss von Fachenthusiasmus auf die Unterrichtsqualität ein Artefakt unspezifischer Messung sein. Folglich sollte zukünftige Forschung untersuchen, inwiefern sich bei kontextspezifischer Messung Enthusiasmus für das Fach Mathematik andere Ergebnisse zeigen.

E-07: Fachliche Lernpfade

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S05 T00 B59
Chair der Sitzung: Constanze Schadl

Arithmetische Fähigkeiten in der Sekundarstufe I als Lernvoraussetzung für die Bruchrechnung digital erheben: Auf dem Weg zu effizienten Kurztests für Lernverlaufsmessungen

Constanze Schadl, Anke Maria Lindmeier
Friedrich-Schiller-Universität Jena, Deutschland

Lernverlaufsmessungen wird insbesondere für formative diagnostische Zwecke ein hohes Potential zugeschrieben (Deno, 2003; Harris et al., 2022). Sie können Lehrkräfte unterstützen, informiertere diagnostische Entscheidungen zu treffen, sodass diese idealerweise den Wissenserwerb von Lernenden gezielter unterstützen können. Damit sind Lernverlaufsmessungen vor allem für solche Lernbereiche von großer Relevanz, welche zum einen einem Großteil der Lernenden Schwierigkeiten bereiten und zum anderen für den weiteren Lernfortschritt zentral sind.

Für den Bereich Mathematik bereitet beispielsweise die Bruchrechnung einem Großteil der Lernenden über den Beginn der Sekundarstufe I hinaus Schwierigkeiten (Vamvakoussi & Vosniadou, 2010; Van Hoof et al., 2015). Zugleich ist dokumentiert, dass die Bruchrechnung für weiteres Mathematiklernen bedeutend ist (Siegler et al., 2012). Längsschnittliche Befunde aus papierbasierten Studien weisen zudem darauf hin, dass verschiedene mathematikspezifische Fähigkeiten als Lernvoraussetzungen Leistungsunterschiede in der Bruchrechnung vorhersagen (Hansen et al., 2015; Schadl, 2020). So werden in der Literatur etwa arithmetische Fähigkeiten zur Multiplikation und Division (symbolisch und in Sachsituationen) als Prädiktoren identifiziert (ebd.). Auch wenn die Zusammenhänge zwischen verschiedenen Prädiktoren und der Bruchrechnung seit kurzem auch differenzierter auf der Grundlage von IRT-skalierten Stufenmodellen beschrieben werden (Schadl, 2020), so können diese Befunde im praktischen Schulkontext von Mathematiklehrkräften noch kaum genutzt werden. Ein Grund dafür ist, dass bisher nur zeitintensive Paper-Pencil-Tests zur Diagnose zur Verfügung stehen, welche in der Schulpraxis unpraktikabel sind.

Die berichtete Studie ist ein erster Beitrag zum übergeordneten Ziel, verfügbare Testinstrumente in ein digitales Setting zu transferieren. Dafür wird der für die Entwicklung digitaler Lernverlaufsmessungen empfohlene dreiphasige Entwicklungsprozess angewendet (Fuchs, 2004), der mit der Prüfung der psychometrischen Qualität des Testinstruments (Phase 1) startet, wobei zudem die empirisch informierte Optimierung der Tests in Bezug auf Testlänge und -passung (effiziente Kurztests) empfohlen wird. Im vorliegenden Beitrag wird die psychometrische Qualitätsprüfung des Testinstruments zu symbolischen Fähigkeiten zur Multiplikation und Division fokussiert. Die Adaption und Digitalisierung des Testinstruments basierte auf papierbasierten Vorarbeiten mit 718 Lernenden der vierten bis sechsten Jahrgangsstufe. Das adaptierte digitale Testinstrument umfasste 22 dichotom kodierte und randomisiert angeordnete Aufgaben zu drei verschiedenen Aufgabentypen (Multiplizieren und Dividieren mit natürlichen Zahlen, Rechenketten mit Platzhaltern, Bestimmung des kleinsten gemeinsamen Vielfachen (kgV)), die aus fachdidaktischer Sicht arithmetische Fähigkeiten als Vorläufer für die Bruchrechnung repräsentieren. Das Testinstrument wurde von 145 Gymnasial- und Regelschullernenden (ca. je zur Hälfte weiblich bzw. männlich) mittels der Onlineplattform Levumi innerhalb von 30 Minuten bearbeitet.

Es zeigten sich eine hohe interne Konsistenz (α = 0.90) sowie auf Basis einer IRT-Analyse gute Itemtrennschärfen und akzeptable Infits zwischen 0.77 und 1.14 (Bond & Fox, 2013). Die Skalenanalysen liefern Hinweise, dass das Aufgabenformat Rechenketten mit Platzhaltern am meisten diagnostischen Informationsgehalt besitzt. Dabei stellen in Einklang mit fachdidaktischen Analysen Rechenketten mit kleinerem Zahlenmaterial sowie mit lediglich einem Multiplikationszeichen niedrigere Anforderungen dar. Rechenketten mit größerem Zahlenmaterial und einem zusätzlichen Divisionsoperator bestimmen dagegen höhere Anforderungen. Die Bestimmung des kgV für theoretisch unterscheidbare Anforderungen zeigt sich hingegen empirisch als wenig differenzierend zwischen verschiedenen Fähigkeitsniveaus. Eine Prüfung der faktoriellen Struktur weist trotz verschiedener Aufgabentypen darauf hin, dass das Testinstrument eindimensional strukturiert ist. Insgesamt deuten die Befunde an, dass das Testinstrument vor dem Hintergrund der Entwicklung von effizienten Kurztests auf das Aufgabenformat der Rechenketten verkürzt werden kann.

Der Kurztest soll in den folgenden Entwicklungsphasen 2 und 3 auf seine Eignung zur Lernverlaufsmessung sowie die praktische Nutzbarkeit durch Lehrkräfte geprüft werden, um gemeinsam mit weiteren in der Entwicklung befindlichen Tests zu Lernvoraussetzungen und Lernergebnissen digitale Lernverlaufsmessungen im Kontext der Bruchrechnung zu ermöglichen.

Entwicklung und Validierung einer Learning Progression zum Erwerb hierarchisch strukturierten Fachwissens

Erika Knack, Vanessa Fischer, Maik Walpuski
Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund

Lernen erfolgt in der Regel kumulativ (Lee, 2012). Ein kumulativer Wissenserwerb erfordert nicht nur eine solide Wissensbasis, sondern auch die Verknüpfung von Wissenselementen (Freiman et al., 2002). Insbesondere in Fächern wie der Chemie, die eine starke hierarchische Struktur aufweisen, ist eine sukzessive Wissensentwicklung essenziell (Fischer et al., 2006). Allerdings belegen verschiedene Studien, dass chemisches Fachwissen von Schülerinnen und Schülern sowie Studierenden sehr heterogen ist (Averbeck, 2021; Behrendt, 2022; Celik & Walpuski, 2018; Hailikari & Nevgi, 2010; Hülsmann, 2015), was – zumindest in Lerngruppen – eine systematische Einbindung neuer Wissenselemente erschwert. Da systematische Wissensstrukturen für kumulatives Lernen unabdingbar sind, ist davon auszugehen, dass Fachwissensdefizite den weiteren Wissenserwerb behindern. Learning Progressions (LP) bieten eine Möglichkeit, eine strukturierte Wissensaneignung zu unterstützen. LP können Lehrkräfte bei der Unterrichtsplanung und Entwicklung von Diagnoseinstrumenten unterstützen, da sie mögliche Lernwege beschreiben, Konzepte (Kernideen) kumulativ über einen bestimmten Zeitraum zu erwerben (Duncan & Hmelo-Silver, 2009; Duschl, Schweingruber & Shouse, 2007). Die hypothetischen Lernwege können über Strand Maps (Beziehungsnetz) dargestellt werden (AAAS, 2007), welche die Kernideen in einer systematisch-hierarchischen Reihenfolge abbilden. Die Abhängigkeiten zwischen den Kernideen werden durch Pfeile dargestellt, die die Richtungen der Lernwege schrittweise anzeigen. Es sind bereits validierte chemie-spezifische LP für den Sekundarbereich I vorhanden (vgl. Celik & Walpuski, 2018; Weber, 2018). Für den Sekundarbereich II wurden bislang keine chemie-spezifischen LP validiert.

Ziele und Forschungsfrage

Ziele des Projekts sind (1) die Entwicklung und fachliche Validierung einer chemie-spezifischen LP für die Sekundarstufe II und (2) die empirische Validierung dieser. Im Rahmen des zweiten Ziels wird folgende Forschungsfrage fokussiert: Kann die angenommene hierarchische Struktur der LP empirisch bestätigt werden?

Studiendesign und Methode

Anhand der Bildungsstandards für die Allgemeine Hochschulreife (KMK, 2020), dem Kernlehrplan für NRW (MSW NRW, 2014) und verschiedenen Schulbüchern wurden chemie-spezifischen Kernideen für die Sekundarstufe II für die verschiedenen Basiskonzepte identifiziert, durch eine Beschreibung von Erwartungen, Grenzen und Schülerfehlvorstellungen charakterisiert und anschließend in einer hierarchisch-logischen Reihenfolge vernetzt. Basiskonzepte sind fachspezifische Erklärungsmuster, die verschiedenen fachlichen Phänomenen zugrunde liegen und so die Grundlage für unterschiedliche Kernideen bilden. Um die LP fachlich zu validieren, wurden die Kernideen sowie die angenommene Vernetzung zwischen diesen von einer Expertengruppe (N=7) inhaltlich validiert und ggf. angepasst.

Für die empirische Validierung wurde zunächst ein Ausschnitt der LP ausgewählt, der sich auf ein Basiskonzept (Chemische Reaktion) bezieht. Zu jeder Kernidee des Basiskonzepts wurden mindestens fünf Items im Multiple-Choice Single-Select-Format entwickelt. Die Testhefte wurden so konstruiert, dass sie immer Items vernetzter Kernideen beinhalten. Diese werden im Rahmen einer Pilotstudie (November 2022) im Multi-Matrix-Design in verschiedenen Oberstufenkursen eingesetzt – wobei 100 Antworten pro Item erwünscht sind. Anschließend werden die Items mittels Item-Response-Theorie analysiert und ggf. überarbeitet. In der Hauptstudie (Februar 2022) werden die überarbeiteten Testhefte erneut im Multi-Matrix-Design in Oberstufenkursen eingesetzt. Anschließend wird auf Grundlage der Daten die hierarchische Struktur der LP analysiert. Mit dem McNemar Test wird der Zusammenhang zwischen zwei Kernideen auf Signifikanz geprüft. Mit Bayes'schen Netzen werden bedingte Zusammenhänge zwischen mehreren Kernideen analysiert.

Ergebnisse

Bisher liegen Daten für die von der Expertengruppe inhaltlich validierte LP vor. Insgesamt wurden 73 für die Oberstufenchemie relevante Kernideen identifiziert – 24 zum Konzept Chemische Reaktion, 14 zum Energiekonzept, 35 zum Konzept vom Aufbau und von den Eigenschaften der Stoffe und ihrer Teilchen. Zwischen diesen Kernideen wurden auf theoretisch validierter Basis 108 Abhängigkeiten ermittelt, wobei eine starke Vernetzung innerhalb der einzelnen Basiskonzepte sowie zwischen dem Konzept Chemische Reaktion und dem Konzept vom Aufbau und von den Eigenschaften der Stoffe und ihrer Teilchen deutlich wird. Die Auswertung des Testinstruments sowie erste Analysen der LP erfolgen im Dezember 2022. Auf der Tagung werden die Methode zur Validierung der LP und erste Ergebnisse hinsichtlich der empirischen Validität der LP präsentiert.

Lernendenschwierigkeiten beim Auswerten von Messdaten – eine Frage des Vorwissens?

Steffen Brockmüller[1], **Mathias Ropohl**[2]
[1]Friedrich-Alexander-Universität Erlangen-Nürnberg; [2]Universität Duisburg-Essen

Theoretischer Hintergrund und Fragestellung

In den Naturwissenschaften sind Lernende beim Auswerten von Versuchsdaten in Denk- und Arbeitsweisen involviert, die das Daten Aufbereiten und Interpretieren, das Schlussfolgern, die kritische Reflexion des Versuchs sowie das Finden möglicher Generalisierungen einschließen. Derartige Operationen fallen Lernenden häufig schwer (z.B. Sandoval & Millwood, 2005). Insbesondere für prozedural komplexere Versuche von Oberstufenlernenden ist allerdings bislang wenig darüber bekannt, welche Schwierigkeiten bestehen. Auch ist nicht geklärt, inwiefern diese Schwierigkeiten mit den dispositionalen Lernvoraussetzungen, insbesondere dem Vorwissen, der Schüler:innen zusammenhängen (vgl. Blömeke et al., 2015). Entsprechend gängiger Konzeptualisierungen des Scientific Reasonings (Opitz et al., 2017), werden fachlich-konzeptuelles, prozedurales (d.h. auf das Auswerten von Versuchsdaten bezogenes) und epistemisches (d.h. auf Konzepte und Begründungen wissenschaftlichen Arbeitens bezogenes) Wissen als mögliche Bedingungsfaktoren vermutet (Osborne, 2018; Schiepe-Tiska, Rönnebeck & Neumann, 2019). Zu klären ist (FF1) welche Schwierigkeiten bei Oberstufenlernenden beim Auswerten chemischer Versuchsdaten vorliegen und (FF2) welche Zusammenhänge sich zwischen dem Auftreten dieser Schwierigkeiten und dem Vorwissen in o.g. Bereichen zeigen.

Methode

Das Vorwissen der Lernenden (N = 154) wurde mithilfe dreier, aus bestehenden Testinstrumenten adaptierten, Skalen (Arnold, 2015; Henke, 2007; Hülsmann, 2015; Nehring, 2014; Vorholzer et al., 2016), erhoben. Analysen der drei Skalen auf Grundlage des Raschmodells zeigen zufriedenstellende Kennwerte hinsichtlich Personenreliabilität (.69 ≤ *PR* ≤ .75), Itemreliabilität (.95 ≤ *IR* ≤ .97) und Itemfit (.47 ≤ Outfit-*MNSQ* ≤ 1.45). Außerdem wurde die Passung von Itemschwierigkeit und Personenfähigkeit mithilfe von Wright-Maps geprüft und für angemessen befunden. Zusätzlich wurde eine nonverbale KFT-Skala eingesetzt. Die Korrelationen der drei Skalen untereinander und zu den kognitiven Fähigkeiten (Heller & Perleth, 2000) fallen erwartungskonform aus und lassen auf diskriminante Validität sowie Nichtkollinearität schließen (.19 ≤ ρ ≤ .63). Anschließend wurde eine Teilstichprobe von n = 72 Schüler:innen beim Durchführen und Auswerten von zwei inhaltlich und strukturell ähnlichen 90-minütigen Experimentiersettings in Zweiergruppen videographiert.

Zur Beantwortung der ersten Forschungsfrage wurden die Videographien mithilfe einer kategorienbildenden qualitativen Inhaltanalyse (Mayring, 2015) induktiv auf Schülerschwierigkeiten untersucht. Die schrittweise Entwicklung eines Kategoriensystems für Schwierigkeiten auf Basis des Videomaterials basierte dabei auf einem angepassten Kodiermanual von Kechel (2016). Zur Beantwortung von Forschungsfrage 2 wurden für diejenigen Schwierigkeitskategorien, für die aufgrund einer ausreichenden Anzahl an Events interpretierbare Ergebnisse zu erwarten waren, binär-logistische Regressionsanalysen des Einflusses der Personenfähigkeiten in den verschiedenen Wissensbereichen auf das Auftreten der Schwierigkeitskategorien durchgeführt.

Ergebnisse und ihre Bedeutung

Die Inhaltsanalyse ergab ein Kategoriensystem, welches 38 Kategorien enthält, die verschiedene Schwierigkeiten beim Auswerten beschreiben. Diese erstrecken sich über alle eingangs

beschriebenen Operationen der Datenauswertung. Genauer liegen 17 Schwierigkeitskategorien für das Daten aufbereiten, je sechs für das Interpretieren, Schlussfolgern und Beurteilen sowie drei für das Verallgemeinern vor. Die logistischen Regressionsanalysen ergeben für einen Großteil dieser Schwierigkeiten zumeist negative signifikante Regressionskoeffizienten – die Wahrscheinlichkeit für das Auftreten steigt also erwartungsgemäß für die meisten Schwierigkeiten mit sinkendem Vorwissen. So steigt beispielsweise die Wahrscheinlichkeit, Messwerte unsystematisch aus Messreihen zum Darstellen auszuwählen, signifikant mit sinkendem epistemischem Wissen. Teils zeigen sich für die beiden Settings unterschiedliche Prädiktoren. Die Auftretenswahrscheinlichkeit der Schwierigkeit „Oberflächliche Interpretation des Graphenverlaufs“ hängt im ersten Experimentiersetting vom prozeduralen, im zweiten Setting vom fachlich-konzeptuellen Wissen ab. Eine Ausnahme bildet die Schwierigkeit „Unsicherheit beim Umgang mit Schwankungen“; hier steigt in Setting 1 die Auftretenswahrscheinlichkeit mit *steigendem* Fachwissen. Offenbar bemerken Schüler:innen mit geringem Fachwissen Unregelmäßigkeiten seltener, sodass diese Schwierigkeit nur für Schüler:innen mit höherem fachlich-konzeptuellem Wissen auftritt.

Die Ergebnisse zeigen, dass Schüler:innen in entsprechenden Lerngelegenheiten nicht ausschließlich auf Fachwissen, sondern auch auf Wissen aus dem Kompetenzbereich Erkenntnisgewinnung zurückgreifen und unterstreichen die Relevanz (insbesondere auch expliziter) Vermittlung entsprechender Inhalte (Petermann & Vorholzer, 2022). Aus methodischer Perspektive bietet die hier erprobte prozessorientierte Untersuchung eine vielversprechende Ergänzung zur bisher üblichen Erhebung mithilfe von Pen-and-Paper-, Simulations- und Hands-on-Testformaten (Schreiber, Theyßen und Schecker, 2014).

Inwiefern führt eine digitale Lernleiter mit niveauadaptiven Hilfen zum Lernerfolg?

Michelle Möhlenkamp[1], Helena van Vorst[1], Sebastian Habig[2], Mathias Ropohl[1]
[1]Universität Duisburg-Essen, Deutschland; [2]Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Theoretischer Hintergrund

Aufgrund der großen Heterogenität der Lernenden an Real- und Gesamtschulen bedarf es einer Differenzierung im Unterricht (Stanat, Schipolowski, Mahler, Weirich & Henschel, 2019). Eine mögliche Methode dafür sind Lernleitern, die individuelle Lernprozesse der Lernenden strukturieren, indem sie ein Thema in mehrere Unterthemen (Meilensteine) einteilen (Girg, Lichtinger & Müller, 2012). Dabei kann eine optimale Passung zwischen dem Lerngegengstand und Lernenden erreicht werden, indem vor und während des Lernprozesses die (Leistungs-)Daten der Lernenden erfasst und für die Bereitstellung niveauadaptiver Hilfen genutzt werden, um so den Wissenszuwachs der Lernenden positiv zu beeinflussen (Brühwiler & Vogt, 2020). Zudem kann der Einsatz digitaler Medien z. B. durch eine automatisierte Erfassung und Auswertung individueller Lernvoraussetzungen zum Erfolg differenzierten Unterrichts beitragen (Hillmayr, Reinhold, Ziernwald & Reiss, 2017). Jedoch liegen bislang nur wenige Studien vor, die die Wirksamkeit des Einsatzes digitaler Medien auf das Fachwissen der Lernenden im Chemieunterricht untersucht haben.

Fragestellung

Basierend auf dem aktuellen Forschungsstand steht folgende Fragestellung im Fokus des Forschungsvorhabens:

Welche Effekte zeigt eine digitale, binnendifferenzierende Lernleiter mit integrierten niveauadaptiven Hilfen auf das Fachwissen, das Interesse und das Selbstkonzept der Lernenden im Chemieunterricht?

Methode

Im Rahmen einer Wissenschaft-Praxis-Kooperation wurde eine digitale Lernleiter mithilfe von H5P zum Thema Atombau für den Chemieunterricht entwickelt. Die digitale Lernleiter besteht aus drei inhaltlichen Meilensteinen (Leitersprossen), die wiederum aus einzelnen Bausteinen zur Strukturierung des Lernprozesses aufgebaut sind. Im gesamten Verlauf der Lernleiter stehen den Lernenden niveauadaptive Hilfen zur Verfügung. In der digitalen Lernleiter erfolgt vor und während des Lernprozesses eine automatisierte Erfassung und Auswertung der Leistungsdaten der Lernenden in Diagnosebausteinen. Basierend auf der Diagnostik werden den Lernenden unterschiedliche, niveauadaptive Hilfen angeboten, z. B. Tippkarten, Zuordnungsaufgaben oder Lückentexte.

Für die Beantwortung der Forschungsfrage wurde eine quasi-experimentelle Interventionsstudie mit Prä-Post-Testdesign durchgeführt, die die Wirkungen einer digitalen Lernleiter auf das Fachwissen und Interesse der Lernenden (Interventionsgruppe) im Vergleich zu einer identischen analogen Lernleiter mit separaten Hilfen erhob (Vergleichsgruppe). 33 Chemieklassen der 9. Jahrgangsstufe aus neun Real- und Gesamtschulen nahmen an dem Projekt teil (N = 697 Lernende).

Ergebnisse und Ausblick

Für die Auswertung konnten N = 323 vollständige Datensätze gewonnen werden. n = 136 Lernende haben mit der analogen, papierbasierten Lernleiter und n = 187 Lernende mit der digitalen Lernleiter gelernt. 52,9 % der Lernenden sind weiblich, 45,5 % männlich und 1,5 % divers. Das durchschnittliche Alter beträgt M = 14.34 Jahre und die letzte Chemie-Zeugnisnote lag im Durchschnitt bei M = 2.78 (befriedigend).

Die Lernenden sind mit gleichen Voraussetzungen gestartet. Im Prätest erzielte die Interventionsgruppe $M_{digital}$ = 10.23 (SD = 3.53), die Vergleichsgruppe M_{analog} = 10.04 (SD = 3.75) Punkte im Fachwissenstest (maximal erreichbare Punktzahl: 37). Eine ANOVA mit Messwiederholung zeigt einen statistisch signifikanten Lernzuwachs für beide Gruppen: $M_{digital}$ = 16.82 (SD = 7.69), M_{analog} = 14.71 (SD = 6.24); $F(1,321) = 233.93$, $p < .001$, $\eta_p^2 = .422$, wobei die Lernenden der Interventionsgruppe einen signifikant höheren Lernzuwachs erzielen ($F(1,321) = 6.85$, $p = .009$, $\eta_p^2 = .021$). Sowohl in der Interventions- als auch in der Vergleichsgruppe haben die Gesamtschüler*innen eine signifikant höhere Punktzahl als die Realschüler*innen im Fachwissenstest erreicht ($F(1,321) = 41.71$, $p < .001$, $\eta_p^2 = .115$).

Eine ANOVA mit Messwiederholung zeigt beim Selbstkonzept für beide Gruppen im Prä-Post-Vergleich einen kleinen Effekt ($F(1,321) = 6.112$, $p = .014$, $\eta_p^2 = .019$). Das Selbstkonzept der Lernenden ist nach der Intervention signifikant höher als vorher. Das individuelle Interesse, die

intrinsische und die extrinsische Motivation haben sich bei beiden Gruppen während der Intervention nicht signifikant verändert.

Die Datenerhebung wurde im September 2022 abgeschlossen, so dass bei der GEBF 2023 die vollständige Analyse der Daten präsentiert wird.

E-08: Berufs- und Karriereentscheidungen

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S05 T00 B71
Chair der Sitzung: Jana Groß Ophoff

Lehramt als Beruf oder Berufung? Berufswahlentscheidungsmotive österreichischer Jugendlicher

Jana Groß Ophoff[1], Gudrun Quenzel[1], Robert Pham Xuan[2], Gernot Brauchle[1]
[1]Pädagogische Hochschule Vorarlberg, Österreich; [2]Universität Innsbruck

Theoretischer Hintergrund

Das Gelingen der im österreichischen Bildungswesen angestrebte Qualitätsentwicklung hängt stark von verfügbaren zeitlichen Ressourcen, der Kompetenz und dem Engagement des pädagogischen Personals ab. Ähnlich wie in vielen anderen Ländern besteht jedoch ein akuter und anhaltender Lehrer:innenmangel an österreichischen Schulen (Huber & Lusnig, 2022), der sich durch den demographischen Wandel in den nächsten Jahren noch verstärken wird (Gehrmann, 2020). Neben Maßnahmen wie die Anstellung von Seiten- und Quereinstieger:innen (Gehrmann, 2019; Porsch, 2021) geraten zunehmend junge Menschen in den Blick, die nach der Hochschulreife vor der Berufswahlentscheidung stehen. Auch wenn mittlerweile eine Vielzahl an empirischen Studien zu der Frage vorliegen, was den Lehrer:innenberuf attraktiv macht (Fray & Gore, 2018), liegt der Fokus v.a. auf (meistens intrinsischen) Berufswahlmotiven (z.B. Pohlmann & Möller, 2010; Watt et al., 2012) operationalisiert werden, anstatt auch Aspekte wie die Übereinstimmung mit beruflichen Interessen, als typisch wahrgenommene Merkmale oder gar „abschreckende“ Aspekte des des Lehrer:innenberufs zu adressieren. Einen entsprechenden Zugang findet man beispielsweise bei Matthes (2019) im Zusammenhang mit der Frage, wie die Attraktivität von Pflegeberufen gesteigert werden kann: Angelehnt an die sog. Zwei-Faktoren-Theorie (Herzberg et al., 2010) unterscheidet die Autorin zwischen Attraktions- und Aversionsfaktoren.

Verbunden mit dem langfristigen Ziel des österreichischen Bundesministeriums für Bildung, Wissenschaft und Forschung, Erkenntnisse für eine Imagekampagne für den Lehrer:innenberuf zu generieren , wird im Herbst 2022 eine repräsentative Befragung unter Maturant:innen an allgemeinbildenden höheren Schulen (AHS) durchgeführt. Im Rahmen dieser Studie werden deren allgemeine schulisch-berufliche Interessen (Bergmann & Eder, 2005) kontrastiert mit parallel formulierten, typischen Attraktions- und Aversionsmerkmalen. Die Auswertung beginnt im Dezember 2022, so dass im Februar erste Ergebnisse vorliegen.

Fragestellung

- Welche schulisch-beruflichen Interessen haben österreichische junge Menschen kurz vor dem Abschluss der Matura?
- Worin unterscheiden sich die beruflichen Interessen von Schüler:innen, die an einem Lehramtsstudium interessiert sind, von jenen, die daran kein Interesse zeigen?
- Welche Attraktivitäts- und Aversionsmerkmale werden von Maturant:innen als typisch für den Lehrer:innenberuf wahrgenommen?
- Inwiefern bestehen zwischen den allgemeinen beruflichen Interessen der Schüler:innen und den typischen Merkmalen des Lehrer:innenberufs eine oder eben keine Passung?
- Welche Schlussfolgerungen lassen sich aus diesen Ergebnissen für eine Erhöhung der Attraktivität des Lehramtsstudiums ziehen?

Methode

Für die im Herbst 2022 durchgeführte repräsentative, Online-Befragung wurde durch das Institut des Bundes für Qualitätssicherung im österreichischen Schulwesen eine nach Bundesland und

Schulgröße stratifizierte Stichprobe österreichischer AHS (n = 109) gezogen. Die Befragung fand im Klassenzimmer statt.

In der Auswertung werden deskriptive sowie multivariate Analysen durchgeführt. Letztere bieten in Form latenter Strukturgleichungsmodelle (z.B. mit Mplus 8, u.a. Muthén & Muthén, 2017) die Möglichkeit, Mess- wie auch Strukturmodelle (unter Berücksichtigung der Mehrebenenstruktur der Daten) zu analysieren sowie die Effekte mehrerer latenter Variablen, aber auch von individuellen Einflussfaktoren gleichzeitig zu modellieren.

Ergebnisse und ihre Bedeutung

Die Datenerhebung läuft derzeit, so dass zur Zeit der Einreichung dieses Abstracts noch keine Ergebnisse präsentiert werden konnten. Im Rahmen des Vortrags sollen erste Befunde zu den oben genannten Fragestellungen vorgestellt sowie weiterführende Analyseperspektiven (wie z.B. Volodina et al., 2015) erörtert werden. Kritisch diskutiert wird, welche Erkenntnisgewinne eine Erweiterung der klassischen Berufsmotivationsstudien um berufsspezifische Attraktions- und Aversionsfaktoren und die Einbeziehung der persönlichen beruflichen Interessen bringen. Eine vielversprechende Referenz zur kritischen Einordnung der in dieser Studie gewonnenen Erkenntnisse stellt die Studie von Eder (2012) dar, in der vor mittlerweile mehr als zehn Jahren die allgemeinen schulisch-beruflichen Interessen österreichischer Schüler:innen am Ende der Pflichtschule im Rahmen von PISA 2009 als österreichische Zusatzanalysen untersucht wurden. Schließlich wird kritisch zu diskutieren sein, wie die Erkenntnisse aus dieser Studie in eine ansprechende Imagekampagne oder andere Maßnahmen zur Steigerung der wahrgenommenen Attraktivität des Lehrer:innnenberufs überführt werden können – v.a. angesichts der Herausforderung einer wissenschaftlichen Gütekriterien genügenden und gleichzeitig bildungsadministrativ-praktisch relevanten Kommunikation zwischen Bildungsforschung und –politik.

Karriereverläufe nach Abbruch des Lehramtsstudiums

Sebastian Franz[1], **Steffen Schindler**[2]
[1]Leibniz Institut für Bildungsverläufe, Germany; [2]Universität Bamberg, Germany

Der Abschluss eines Lehramtsstudiums ist ein bedeutender Schritt auf dem Weg zum Lehrkraftberuf in Deutschland. Allerdings schließen nicht alle Studierende, die sich für ein Lehramtsstudium einschreiben, das Studium mit einem Abschluss ab. Im Vergleich zu anderen Studiengängen ist die Anzahl der Abbrüche im Lehramtsstudium eher niedrig. Während die durchschnittliche Abbruchrate an deutschen Hochschulen bei 31 % liegt, brechen nur 21 % der Lehramtsstudierenden ihr Studium ab (Heublein et al., 2022). Trotz der niedrigen Abbruchquote stellt die Nichtbeendigung des Lehramtsstudiums ein bildungspolitisches Problem dar, da Deutschland aktuell mit einem massiven Lehrermangel konfrontiert ist. Eine Möglichkeit, dieser Entwicklung entgegenzuwirken, ist es, mehr Lehramtsstudierende zu einem Abschluss zu führen.

Vor diesem Hintergrund ist es von entscheidender Bedeutung zu verstehen, warum Studierende ihr Lehramtsstudium trotz klar definiertem Berufsweg mit hervorragenden Arbeitsmarktaussichten abbrechen. Generell kann man zwei Hauptgründe für den Studienabbruch unterscheiden: Einerseits können Studierende die akademischen Anforderungen der Studiengänge nicht erfüllen. Andererseits

können sich die Studierenden für einen anderen Berufsweg entscheiden, der attraktiver erscheint als eine Karriere als Lehrkraft (Herfter et al., 2015). In diesem Beitrag konzentrieren wir uns auf Letzteres und beschäftigen uns explizit mit den Karrierewegen, die Alternativen zur ursprünglichen Entscheidung für den Lehrkraftberuf darstellen. Solche Alternativen können der vollständige Abbruch des Hochschulstudiums oder der Wechsel zu einem anderen Studiengang sein. Eine Besonderheit des Lehramtsstudiums besteht darin, dass die Studierenden ihre Studienfächer beibehalten, aber einen anderen Studiengang anstreben können, der in andere Berufe als den der Lehrkraft führt.

Fragestellung und theoretischer Rahmen

Ziel des Beitrags ist es, typische Muster des Studienabbruchs in Lehramtsstudiengängen zu identifizieren. Bei diesem Ansatz muss berücksichtigt werden, dass es aus individueller Sicht rational sein kann, ein Lehramtsstudium nicht abzuschließen, weil andere Alternativen attraktiver erscheinen. Daher sollen die Faktoren untersucht werden, welche für solche Verschiebungen in der Wahrnehmung der Attraktivität von Studiengängen verantwortlich sind. Den theoretischen Rahmen zur Untersuchung dieser Faktoren bilden soziologische Entscheidungsmodelle zu Bildungs- und Karriereentscheidungen aus der Familie der Rational-Choice-Theorie (Breen & Goldthorpe, 1997), die wir hier auf die Revidierung der ursprünglichen Entscheidung für den Lehramtsberuf anwenden. Demnach sollte die gewählte Alternative das Nutzen-Kosten-Kalkül des Lehramtsstudiums übersteigen und eine höhere Aussicht auf Erfolg bieten.

Auf diese Weise sollen Erkenntnisse gewonnen werden, die als Grundlage für politische Maßnahmen zur Steigerung der Attraktivität von Lehramtsstudiengängen dienen können.

Daten und Methoden

Die Datenbasis der vorliegenden Untersuchung ist eine Stichprobe von Lehramtsstudierenden aus der Startkohorte 5 des Nationalen Bildungspanels (NEPS; Blossfeld & Roßbach, 2019), die sich über alle in Deutschland angebotenen Lehramtstypen und Fächergruppen erstreckt. Das Sample enthält Karriereverläufe von 5500 Lehramtsstudierenden, wovon 350 ihr Lehramtsstudium nachweislich abgebrochen oder nach erfolgreichem Bachelor of Education keinen Master of Education begonnen haben.

Als statistische Methode wird die Sequenzmusteranalyse angewendet, die auf einer Kombination von Optimal Matching und Clusteranalyse basiert und einen Vergleich zwischen Lebensläufe einzelner Individuen ermöglicht (Sackmann, 2013). Die Statusangaben wurden monatsgenau aufbereitet.

Ergebnisse

Erste Befunde aus den Voranalysen zeigen, dass ca. 60 % der Befragten einen Monat nach Austritt aus dem Lehramtsstudium immer noch an einer Hochschule studierten. Lediglich 8 % haben eine Ausbildung begonnen und 16 % gingen einer regulären Erwerbstätigkeit nach. Ungefähr 12 % befinden sich in einem Praktikum, Elternzeit oder Wehrdienst bzw. freiwilligen Dienst. Nur ein sehr geringer Anteil ist mit 4 % erwerbslos.

In der geplanten Präsentation werden die identifizierten Verlaufsmuster dargestellt und hinsichtlich ihrer Implikationen diskutiert. Dabei soll der Fokus auf der Erklärung liegen, warum

Lehramtsstudierende andere Karrierewege als attraktiver ansehen als andere und aus dem Lehramtsstudium austreten.

Promovieren mit Lehramtshintergrund – Monitoring und Unterstützung

Regina Bedersdorfer, Hendrik Lohse-Bossenz
Pädagogische Hochschule Heidelberg, Germany

Theoretischer Hintergrund

Ziel dieses Beitrages ist ein differenziertes Monitoring des Promotionsprozesses und damit eine gezielte Unterstützung von Promovierenden mit Lehramtshintergrund. Die wissenschaftliche Weiterqualifizierung von (angehenden) Lehrkräften via Promotion ist essentiell um Schul- und Unterrichtsforschung weiterzuentwickeln (KMK, 2014). Dabei ist die Promotionsphase ein Lernprozess mit vielfältigen Herausforderungen (Fiedler & Hebecker, 2012). Ein differenziertes Monitoring dieses Prozesses kann helfen, Probleme frühzeitig zu analysieren und angemessene Unterstützungsformen zu finden. Nach Vogel, Brandt, and Jaksztat (2017) muss eine angemessene Unterstützung den Bedürfnissen der Promovierenden entsprechen.

Hierbei haben Promovierende mit Lehramtshintergrund besondere Unterstützungsbedarfe. Zusätzlich zu allgemeinen Promotionsherausforderungen (z. B. Infragestellen der Relevanz eigener Forschung) (Fiedler & Hebecker, 2012), werden sie mit Hürden konfrontiert. Hierzu zählen die Sozialisation in das wissenschaftliche Umfeld (Bakx, Bakker, Koopman, & Beijaard, 2016) oder die geringe forschungsmethodische Ausbildung im Lehramtsstudium (Stelter, 2019). Erste quantitative Ergebnisse weisen auf einen vergleichsweise hohen Zusammenhang zwischen hoher emotionaler, fachlicher sowie netzwerkintegratorischer Unterstützung und Promotionsfortschritt von Individuen mit Lehramtshintergrund hin. Zudem verändern sich Unterstützungsbedarfe im Promotionsverlauf (Bedersdorfer & Lohse-Bossenz, 2022).

Forschungsfrage

Demgemäß fragt dieser Beitrag danach, wie ein Verfahren aussehen kann, welches den Promotionsprozess von Individuen mit Lehramtshintergrund differenziert nachverfolgen und zur Analyse angemessener Unterstützunformen in unterschiedlichen Promotionsstadien beitragen kann? Hierfür wird zunächst ein Index für die Erfassung des Fortschritts von Dissertationen – bestehend aus elf Teilaufgaben – in den Bildungswissenschaften anlehnend an Döring and Bortz (2016) entwickelt und validiert. Zudem wird die Bedeutung der Unterstützungsarten *emotional, fachlich* und *Netzwerkintegration* aus dem Lernumweltmodell nach Vogel et al. (2017) für einen hohen Promotionsfortschritt analysiert.

Methodik

Zur ersten Validierung des Index wurden im Oktober 2018 alle Promovierenden einer Pädagogischen Hochschule (N = 101, Rücklauf 52,5%) schriftlich befragt. Sie sollten zunächst Auskunft über den Gesamtstand ihrer Promotion (M = 54,19%, SD = 31,09%) und ihre Unterstützungssituation geben. Anschließend wurde eine Selbsteinschätzung des Fortschrittes innerhalb elf Teilaufgaben, jeweils auf einer Skala von 0-100%, abgefragt. Der aggregierte Promotionsstand (M = 62,85%, SD = 27,03%) berechnet sich aus dem Mittelwert der Prozentangaben zu den Teilaufgaben. Mittels multivariater

OLS soll geprüft werden, inwiefern die Angaben der Promovierenden zur ihrem Gesamtstand (AV) mit dem aggregierten Promotionsstand (UV) übereinstimmen.

16 Personen aus dieser Gruppe gaben ein Jahr später erneut Auskunft zu ihrer Promotion. So kann der Fortschritt innerhalb eines Jahres betrachtet und analysiert werden. Dabei verfolgt dieser Beitrag einen fallorientierten Zugang. Er nutzt für die Analyse der Bedeutung von verschiedenen Unterstützungsformen in unterschiedlichen Promotionsstadien (Bedingungen) für einen hohen Fortschritt in den einzelnen Promotionsschritten (Outcome) die Qualitative Comparative Analysis (QCA) in einem Crisp-Set (0-1 Variablen-Kodierung).

Ergebnisse

Die Validierung des Index mittel OLS zeigt, dass der aggregierte Promotionsstand die Einschätzung zum Gesamtstand unter Berücksichtigung verschiedener Dissertations- und Forschungsarten gut widerspiegelt (N = 50; B = .915; p < .01). Dies spricht für eine zuverlässige Selbsteinschätzung der Promovierenden.

Die zweite Befragung von 16 Lehramtspromovierenden zeigt weiterhin, dass diese innerhalb eines Jahres ihren Fortschritt sehr unterschiedlich einschätzen (Fortschritt: M = 18.60, Min = -1.82, Max = 37.27) und die Betrachtung der einzelnen Teilaufgaben einen differenzierten Einblick in individuelle Promotionsvorhaben ermöglicht.

Die Ergebnisse der QCA veranschaulichen ferner, dass für einen hohen Fortschritt der Teilaufgaben verschiedene Arten hoher Unterstützung in verschiedenen Promotionsstadien hinreichend sind. Die Befragten profitieren beispielsweise bei der Aufarbeitung des Forschungsstandes von hoher emotionaler und netzwerkintegratorischer Unterstützung zu Beginn der Promotion. Fachliche Unterstützung ist im Besonderen bei der Einordnung der Forschungsergebnisse im fortgeschrittenen Stadium der Promotion hilfreich.

Die Ergebnisse zeigen, dass ein differenziertes Monitoring des Promotionsprozesses dazu beitragen kann, Probleme von Promovierenden mit Lehramtshintergrund frühzeitig aufzudecken und zu analysieren. Zudem kann die Bedeutung einzelner Unterstützungsmaßnahmen gezielter analysiert werden. Dies ist besonders für Betreuende und verantwortliche Einrichtungen relevant – schließlich sollen diese Promovierende angemessen unterstützen (Wissenschaftsrat, 2011).

Konsequenzen der Kompromissbildung im Übergang in Berufsausbildung für die Ausbildungsstabilität

Brigitte Schels[1], Corinna Kleinert[2], Lea Ahrens[2], Melanie Fischer-Browne[2]
[1]Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany; [2]Leibniz-Institut für Bildungsverläufe

Viele Auszubildende in Deutschland machen eine Berufsausbildung in einem anderen Beruf, als noch in der Schulzeit angestrebt. In dieser Studie betrachten wir die Konsequenzen solcher beruflichen Kompromisse für die Stabilität der Berufsausbildung. Dies ist ein wichtiges Thema, da in Deutschland im Schnitt jede vierte Berufsausbildung, hier betriebliche Ausbildung, vorzeitig beendet wird.

Berufswahltheorien spezifizieren, dass nach Eintritt in Ausbildung der Ausbildungsberuf auf den Prüfstand gestellt und die Berufswahl bei erlebter Nicht-Passung ggf. überdacht, eine Ausbildung nicht weitergeführt wird (Herzog et al. 2006). Für die Frage der Nicht-Passung können

unterschiedliche Berufsdimensionen eine Rolle spielen. Nach der Berufstheorie von Gottfredson (2002), die auch den Begriff der Kompromissbildung geprägt hat, streben junge Menschen nach Berufen, deren Geschlechtstypik, Prestige und Interessensprofil im Einklang mit dem Selbst stehen. Dabei argumentiert Gottfredson, dass Jugendliche in der Formation von realistischen Berufsaspirationen, d.h. den an die wahrgenommenen Opportunitätsstrukturen angepassten Aspirationen, eher Einschnitte in den Interessen in Kauf nehmen als in der Geschlechtstypik oder im Prestige. Sobald die Entscheidung für eine erste Berufsausbildung gefallen ist, legt dagegen die Person-Environment-Fit-Theorie (Edwards & Shipp, 2009) nahe, dass die Jugendlichen erstmals die Passung von Tätigkeit und angestrebten Berufsinteressen erfahren und dies daher besonders bedeutsam wird. Wir nehmen daher an, dass Interessenskompromisse die Ausbildungsstabilität stärker gefährden als Kompromisse in Prestige und Geschlechtstypik. Schließlich bestehen in der Berufswahl Unterschiede im Geschlecht. So können sich die Konsequenzen in der Berufswahl zwischen Männern und Frauen unterscheiden, wenn Kompromisse in Geschlechtstypik und Prestige bei jungen Männern eher als Abweichung von den sozialen Erwartungen erlebt werden als bei Frauen. Wir nehmen daher an, dass sich Kompromisse in Form von weniger geschlechtstypischen oder weniger prestigeträchtigen Berufen bei Männern stärker auswirken als bei Frauen.

In der Studie betrachten wir, wie sich Kompromisse zwischen Ausbildungsberuf und realistischen Berufsaspiration auf das Risiko einer vorzeitigen Beendigung der Berufsausbildung auswirken: Wie bedeutsam sind Kompromisse in Geschlechtstypik, Prestige und Interesse? Gibt es dabei Unterschiede zwischen jungen Männern und Frauen?

Die Analyse basiert auf einer Stichprobe von mehr als 3000 Auszubildenden aus der Kohorte von Neuntklässlern, die im Jahr 2010 im Nationalen Bildungspanel (NEPS-SC4) befragt wurden. Die Befragungsdaten wurden mit aggregierten Berufsdaten der Magnitude Prestige Skala, zum Frauenanteil in der Berufen (Mikrozensus) und Interessensprofilen (O*Net) kombiniert. Kompromisse werden als Differenz in jedem der drei Merkmale zwischen Ausbildungsberuf und dem in der Klasse 9 erhobenen realistischen Berufsaspirationen operationalisiert. Das Risiko einer vorzeitigen Beendigung einer Berufsausbildung, hier der ersten betrieblichen Ausbildung oder Ausbildung an einer Berufsfachschule, wird mit diskreten Hazard-Rate-Modellen geschätzt, in denen die Kompromissindikatoren als Prädiktoren und weitere Kontrollvariablen eingehen.

Die Befunde zeigen, dass Jugendliche eher ihre Ausbildung vorzeitig beenden, wenn ihr Ausbildungsberuf weniger geschlechtstypisch ist als ihre Berufsaspiration aus der Schulzeit. Dieser Zusammenhang besteht jedoch nicht unabhängig von Interessenskompromissen. In einer gemeinsamen Analyse gefährden insbesondere die Interessenkompromisse die Stabilität der Berufsausbildung. Die Effekte von Kompromissen in der Geschlechtstypik gehen dabei vollständig in den Effekten des Interessenskompromisses auf. Während es keine Geschlechterunterschiede in den Konsequenzen von Interessenskompromissen gibt, zeigt sich ein höheres vorzeitiges Beendigungsrisiko bei den jungen Männern, deren Ausbildungsberuf weniger geschlechtstypisch ist als ihr angestrebter Beruf. Nur bei den jungen Frauen spielen Kompromisse im Prestige eine Rolle, wirken aber gar ausbildungsstabilisierend.

Die Studie zeigt, dass eine Stabilität der Berufsausbildung gefährdet ist, wenn Jugendliche in Ausbildungsberufe eintreten, die nicht im Einklang mit den in der Schulzeit gebildeten Interessensprofilen ihrer Berufsaspirationen stehen. So ist eine interessensorientierte Berufsberatung und -orientierung auch in Zeiten von Passungsproblemen am Ausbildungsmarkt nicht zu vernachlässigen. Dabei gehen Interessenskompromisse oft Hand in Hand mit Kompromissen in der Geschlechtstypik. Die verbleibenden Unterschiede in den Effekten einer geschlechtsuntypischeren Berufswahl für die Ausbildungsstabilität zwischen jungen Männern und Frauen zeigen, wie „doing gender“ auch in der Berufswahl eine Rolle spielen.

E-09: Beraten und Beurteilen

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S05 T00 B83
Chair der Sitzung: Philipp Wotschel

Beratungskompetent durchs Lehramtsstudium? Entwicklung und Erprobung eines handlungsnahen Prüfungsformates

Philipp Wotschel
Universität Paderborn, Deutschland

Beratung bildet eine zentrale Tätigkeit von Lehrkräften. Seit den Erkenntnissen der ersten PISA-Studie werden Lehrkräfte seitens der Elternschaft vermehrt adressiert, um Aufschluss über die Qualität der Ausbildung ihrer Schüler*innen zu erhalten (Gartmeier, 2018). Zugleich sehen sich Lehrkräfte einer heterogeneren Schülerschaft mit unterschiedlichen Lebensrealitäten gegenüber, deren Beratungsanliegen zunehmen (Schnebel, 2007). Zwar ist der Erwerb von Beratungskompetenzen als KMK-Standard (2019) formuliert und an vielen Hochschulen curricular verankert (bspw. Modulhandbuch M.Ed. Universität Paderborn 2016), dennoch sieht sich ein erheblicher Teil der Lehrkräfte durch ihre akademische Ausbildung nicht ausreichend auf künftige Beratungsarbeit vorbereitet (Hertel & Schmitz, 2010). In Anlehnung an andere professionsorientierte Ausbildungen, wie der Medizin, bieten rollenspielbasierte Gesprächssimulationen das Potenzial, die Beratungskompetenz (angehender) Lehrkräfte zu fördern (Dotger et al., 2010; Gerich, 2016), da hierbei typische Standardanforderungen des Berufs unter möglichst authentischen, aber komplexitätsreduzierten Bedingungen simuliert werden (St.Pierre & Breuer, 2013). Derartige simulationsbasierte Settings fokussieren den handlungsnahen, performativen Anteil professioneller Kompetenz (Blömeke et al., 2015), der im Lehramtsstudium typischerweise nicht Gegenstand von Prüfungen ist. So sind die Kompetenzen, die in der ersten Phase der Lehramtsausbildung, etwa durch Klausuren oder mündliche Prüfungen, geprüft werden, eher auf die Erfassung von Wissen und nicht von handlungsnahen Fähigkeiten ausgerichtet (Miller, 1990). Die Prüfung handlungsnaher Kompetenzen kommt erst in der zweiten Phase der Lehramtsausbildung in unterrichtspraktischen Prüfungen zum Tragen, die jedoch sofort einen Umgang mit der vollen Komplexität des Klassenunterrichts erfordern. Es zeichnet sich im Lehramtsstudium also eine Lücke zwischen handlungsfernen Prüfungen im Studium und Prüfungen unter unstandardisierten Bedingungen des realen Berufsfeldes ab, die sich auch darin äußert, dass es Studierenden oft unzureichend gelingt, ihr erworbenes Wissen in solchen Handlungskontexten adäquat anzuwenden (Zeichner, 2010). Der Einsatz von handlungsnahen Prüfungs- und Lehrformaten beziehungsweise Performanztests könnte einen Beitrag dazu leisten, diese Lücke zu verkleinern.

Als möglicher Prototyp einer standardisierten, handlungsnahen Prüfung wurde ein Performanztest mit zugehörigen Bewertungsmodell zur Erfassung und Beurteilung beratungsbezogenen Verhaltens Lehramtsstudierender für die Sekundarstufe I in simulierten Beratungsgesprächen entwickelt. Trainierte Akteur*innen interagieren in Simulationsgesprächen als Elternteile und Schüler*innen mit Studierenden als Lehrkräften in einem zwei-stufigen Rollenspiel, in denen über lern- und sozialverhaltensbedingte Aspekte eines Schulkindes beraten werden soll (Klassenstufe 9). Basierend auf Befunden zu schulischen Beratungsgesprächen (Wegner, 2016) wurde dabei eine prototypische Beratungssituation entwickelt und hinsichtlich Repräsentativität und Authentizität durch Expertenbefragungen validiert. Die Ausbildung der Akteur*innen zur Einübung der Rollen und zum

Umgang mit so genannten „verbal triggers“ zur Standardisierung der Prüfungssituation (Dotger et al., 2010) erfolgte in einem achtwöchigen Trainingszeitraum. Im Rahmen einer Pilotuntersuchung wurden mit teilnehmenden Studierenden stimulated-recall-Interviews zur weiteren Validierung der „verbal trigger“ und leitfadengestützte Rückmeldeinterviews zur Simulationswahrnehmung und -akzeptanz durchgeführt und strukturierend-inhaltsanalytisch ausgewertet (Kuckartz, 2018). Die derzeit vorliegenden Ergebnisse (September 2022, N=8) weisen auf eine gute Passung zwischen den konzeptuellen Vorüberlegungen und der Durchführbarkeit der Gesprächssimulationen hin, insbesondere dass der Einsatz von Hilfskräften für performanzorientierte Prüfungen realisierbar ist. Insgesamt zeichnet sich ab, dass „verbal trigger“ bestimmte intendierte Beratungskompetenzfacetten, wie soziale Kooperationskompetenz oder pädagogisches Wissen (Hertel, 2009), evozieren können und dabei teilweise auch nicht intendiert die Dimension der Bewältigungskompetenz (Bruder, 2011) triggern. Die Studierenden nehmen die simulierten Gespräche und die Akteur*innen als authentisch war, was unter anderem durch den Handlungsdruck in den Simulationen unterstützt wird. Der Einsatz von simulierten Beratungsgesprächen wird dabei nach der Simulationsteilnahme als mögliches Prüfungsformat akzeptiert und im Vergleich zu schriftlichen und mündlichen Prüfungen als authentischer, praxisrelevanter und objektiv beschrieben. Dies jedoch unter der Bedingung, dass eine entsprechende Vorbereitung auf ein solches Format, beispielsweise im Rahmen einer Lehrveranstaltung, durchgeführt wird. In der anschließenden Hauptuntersuchung (Wintersemester 2022/2023) wird der entwickelte Performanztest in einer Querschnittserhebung mit Lehramtsstudierenden eingesetzt, um einen quantitativen Score zur Bewertung von Beratungsperformanz auf breiterer empirischer Basis zu fundieren und Zusammenhänge zu weiteren Facetten professioneller Kompetenz (z.B. Beratungswissen) zu untersuchen.

Beratungskompetenz von Grundschullehrkräften bei hoher Begabung – eine Modellentwicklung

Sonja Seiderer, Stefanie Schnebel, Robert Grassinger
Pädagogische Hochschule Weingarten, Deutschland

Theoretischer Hintergrund

Beratung von Schülerinnen und Schülern stellt eine wesentliche und verpflichtende Aufgabe von Lehrkräften dar (u.a. Kultusministerium Baden-Württemberg, 2000; Sauer, 2019; Schnebel, 2017). Die Beratungsanlässe hierbei sind vielfältig und weisen ein großes inhaltliches Spektrum auf (u.a. Gerich et al., 2014; Grewe, 2015; Hertel & Schmitz, 2010). Um diesen variierenden Beratungssituationen professionell begegnen zu können, benötigen Lehrkräfte Beratungskompetenzen, die in verschiedenen allgemeinem Beratungskompetenzmodellen modelliert sind (Bülow, 2018; Bruder et al., 2010; Gerich et al., 2015; Hertel, 2009). Strasser (2020) postuliert, dass die verschiedenen Beratungsanlässe spezifische Kompetenzfacetten erfordern. Auch in der Kompetenzforschung zeigt sich, dass Kompetenzen mehrdimensional und situationsspezifisch zu betrachten sind (Blömeke et al. 2016) und situationsspezifische Handlungssituationen spezifisches Wissen und Fähigkeiten erfordern (u.a. Buiskool & Broek, 2012).

Als Beispiel für eine spezifische Beratungssituation für Grundschullehrkräfte wird hier die schulische Beratung im Kontext hoher Begabung betrachtet (Müller-Oppliger, 2021; Nguyen & Sliwka, 2021).

Dabei werden Beratungsanlässe hinsichtlich begabungsspezifischer Fördermaßnahmen wie die Teilnahme an Enrichmentprogrammen (z. B. Kinderakademien) oder Akzelerationsmaßnahmen (z. B. das Überspringen einer Klasse) in den Blick genommen. Diesbezüglich wurde unter Berücksichtigung von Forschungsergebnissen (1) zum Wissen über Begabung (Heyder et al., 2018), (2) zu Stereotypen von Lehrkräften bezüglich Begabter (Baudson & Preckel, 2016), (3) zur Identifizierung und Förderung begabter Kinder (Bergold, 2011; Nguyen & Sliwka, 2021) und (4) zu grundlegenden Modellen von Beratungskompetenz ein spezifisches, mehrdimensionales Kompetenzmodell entwickelt. Konkret werden für die Beratungskompetenz von Grundschullehrkräften bei hoher Begabung die Dimensionen Wissen, Einstellungen und Überzeugungen, Diagnostische Kompetenz und Beratungskompetenz als relevant postuliert. Die strukturelle Prüfung dieses Modells sowie die Untersuchung dessen Validität ist Ziel der Studie.

Methodik

Präsentiert werden Ergebnisse aus Online-Befragungen von Lehrkräften, Schul- psycholog*innen und Berater*innen externer Beratungsstellen für hohe Begabung. Die Stichprobe umfasst 156 Teilnehmende (*MAlter*=44.8;*SD*=9,5), davon 112 Lehrkräfte, 33 Schulpsycholog*innen und 15 Berater*innen externer Beratungsstellen. Die Dimension Wissen wurde über einen Wissenstest erhoben, welcher im Rahmen eines nominalen IRT-Modells (Bock, 1972) ausgewertet wurde. Die Überprüfung der Dimension Überzeugungen / Einstellungen erfolgte anhand eines semantischen Differentials, ausgewertet über (a) die Bildung von Scores, (b) eine latente Klassenanalyse und (c) einer persönlichen Einschätzung zur Förderung bei hoher Begabung und Beratung. Die Dimensionen Diagnostische Kompetenz und Beratungskompetenz schließlich wurden mittels eines Situational Judgment Tests (SJT) in Anlehnung an Bruder et al. (2017) operationalisiert.

Ergebnisse

Die Befunde verweisen auf eine akzeptable bis gute Reliabilität der einzelnen Skalen für die vier Dimensionen des Modells (Wissen: *EAP*= .65 – 90; Überzeugungen*:* = .73 - .84; Diagnostische Kompetenz und Beratungskompetenz: *ICC*= .80 - .95). Die angenommene Struktur zeigte zunächst unzufriedenstellende Modellfits, insbesondere in Bezug auf den CFI-Wert (^{2}=151.54; *df*=141; (*p*=.24); *CFI* =.90; *RMSEA*= .03; *SRMR*= .08). Durch die Betrachtung des Perspektivwechsels – ursprünglich als Aspekt Diagnostischer Kompetenz modelliert (Gerich et al., 2014) – als Extra-Dimension ergab sich jedoch durchgängig ein zufriedenstellender Modellfit. (^{2}=137.54; *df*=133; (*p*=.38); *CFI* =.962; *RMSEA*= .02; *SRMR*= .07). Außerdem gaben erwartungskonforme, signifikante Mittelwertsunterschiede zwischen Lehrkräften und Schulpsycholog*innen bzw. Mitarbeitenden externer Beratungsstellen von mittlerer bis großer Effektstärke Hinweise auf die Konstruktvalidität (*d*= .37 - 1.46).

Diskussion

Die Ergebnisse bestärken für die Beratung begabter Schülerinnen und Schüler die Annahme von Strasser (2020), die Beratungskompetenz anlassspezifisch zu betrachten. Hierbei scheint entgegen des Modells von Gerich et al. (2014) der Perspektivwechsel nicht einer Diagnostischen Kompetenz immanent zu sein, sondern als eigene Kompetenzfacette betrachtet werden zu müssen. Der Hinweis

auf die Existenz eines situationsspezifischen Modells für die Beratungskompetenz bei hoher Begabung wirft die Frage auf, inwieweit auch für weitere Beratungsanlässe, die Lehrkräften im schulischen Alltag begegnen, anlassspezifische Beratungskompetenzmodelle anzunehmen sind.

Verringern irrelevante Informationen die Qualität von Lehrkräfteurteilen über die Motivation von Schüler:innen?

Jan Beck, Stephan Dutke, Till Utesch
Westfälische Wilhelms-Universität Münster, Deutschland

Motivation von Schüler:innen zu diagnostizieren, zählt zu den zentralen, professionellen Kompetenzen von Lehrpersonen. Allerdings stimmen informelle Urteile von Lehrpersonen oft nur in geringem Maße mit den Selbsteinschätzungen der Schüler:innen überein (Urhahne & Wijnia, 2021). Zielorientierungen (Elliot & Church, 1997) scheinen für die Motivation von Schüler:innen relevant zu sein, jedoch berücksichtigen Lehrpersonen vermehrt nicht-diagnostischer Informationen (z. B. Geschlecht oder Leistung; Brandmiller et al., 2020; Kaiser et al., 2013). Daneben vernachlässigen Lehrpersonen diagnostische Informationen (Dicke et al., 2012). Das Vorhandensein nicht-diagnostischer Informationen könnte die Verwendung diagnostischer Informationen beeinträchtigen. Hieraus ergibt sich folgende Fragestellung:

Wie unterscheidet sich die Güte des Motivationsurteils, wenn nur diagnostische (Bedingung 1) bzw. diagnostische und nicht-diagnostische Informationen (Bedingung 2) vorhanden sind?

Zur Untersuchung dieser Frage wurde eine Online-Vignettenstudie durchgeführt. N = 105 (M_{Alter} = 34.7 Jahre, SD_{Alter} = 12.7) Lehramtsstudierende (n = 45) und Lehrpersonen (n = 60) wurden zufällig zwei Bedingungen zugewiesen, in denen jeweils acht fiktive Schüler:innen hinsichtlich ihrer Motivation beurteilt wurden. Die Personen in Bedingung 1 (n = 63) erhielten Vignetten mit ausschließlich diagnostischen Informationen (Lernzielorientierung und Arbeitsvermeidungstendenzen; jeweils „niedrig“, „durchschnittlich“, oder „hoch“). Personen in Bedingung 2 (n = 42) erhielten die gleichen Vignetten ergänzt um Informationen über Geschlecht (m, w) und Leistung (unter- oder überdurchschnittlich), die nicht-diagnostisch für Motivation sind. Die daraus resultierenden 32 Vignetten wurden in 4 Materialsets (n_1 = 13, n_2 = 13, n_3 = 9, n_4 = 7) von je 8 Vignetten gegliedert. Jede Person in Bedingung 2 erhielt eines dieser Sets, so dass alle Kombinationen diagnostischer und nicht-diagnostischer Informationen gleich häufig bearbeitet wurden.

Um Motivationsurteile über fiktive Schüler:innen, solchen über reale Schüler:innen anzunähern, sollten Versuchspersonen heterogene Profile fiktiver Schüler:innen für eine Projektwoche in Gruppen einteilen. Schüler:inneninformationen wurden somit ohne Wissen über die diagnostische Aufgabe präsentiert, jedoch mit dem Ziel, Informationen zu behalten.

In beiden Bedingungen hängen die Motivationseinschätzungen unterschiedlich stark mit Lernzielen (Bedingung 1: η^2 = .223; Bedingung 2: η^2 = .021) und Arbeitsvermeidung (Bedingung 1: η^2 = .017; Bedingung 2: η^2 = .019) zusammen. Der Unterschied zwischen den Bedingungen ist in der Nutzung von Lernzielen signifikant (β = -.28; p < .001), von Arbeitsvermeidung nicht (β = -.08; p = .255). In Bedingung 2 sind Motivationsurteile zusätzlich unterschiedlich zwischen über- und unterdurchschnittlich leistungsstarken Schüler:innen (η^2 = .140), aber nicht zwischen den Geschlechtern (η^2 = .028).

Die Ergebnisse zeigen, dass diagnostische Informationen von Motivation genau dann weniger genutzt werden, wenn zusätzlich Informationen über die Leistung und das Geschlecht der Schüler:innen vorhanden sind. Daraus folgt die Vermutung, Lehrpersonen priorisieren Leistungscharakteristika in alltäglichen Schulsituationen im Gegensatz zu den motivational relevanten Hinweisen. Für retrospektive Urteile über die Motivation könnten somit Leistungscharakteristika eine für Lehrpersonen gewichtige Information sein statt Lernziele bzw. Arbeitsvermeidung. Folglich sollte diskutiert werden, inwiefern der Stellenwert von Leistung in der Schule die Wahrnehmung der Lernbereitschaft von Schüler:innen verzerrt.

Wenig kompetentes Sozialverhalten als Risikofaktor? Lehrkraftemotionen, Communion und Agency des Verhaltens der Lehrkraft in Abhängigkeit des Sozialverhaltens des Lernenden

Karoline Koeppen, Madeleine Kreutzmann, Madita Frühauf, Bettina Hannover
FU Berlin, Deutschland

Schüler:innen profitieren von einer qualitätsvollen Beziehung zu ihrer Lehrkraft (Metaanalysen: Roorda et al., 2011; Roorda et al., 2017) und von positiven Emotionen der Lehrkraft (z.B. Frenzel et al., 2021) in ihrer Motivation und ihren Leistungen. In einer qualitätsvollen Beziehung zeigt die Lehrkraft warmes Verhalten, wendet sich dem Lernenden zu und reagiert feinfühlig auf seine Bedürfnisse (hohe Communion). Gleichzeitig motiviert und fördert eine qualitätsvolle Beziehung den Lernenden, weil die Lehrkraft hier das Ausmaß, in dem sie das Verhalten des Lernenden steuert und überwacht (Agency) davon abhängig macht, wie stark die Kompetenzen des Lernenden dies erfordern (Hannover et al., 2022). Positive Lehrkraftemotionen haben einen günstigen Effekt auf den Lernenden, weil sie die Qualität der Lehrkraft-Schüler:in und die Unterrichtsqualität steigern, aber auch, weil sie sich direkt über emotionale Ansteckung auf den Lernenden übertragen (Frenzel et al., 2021).

Dem Komplementaritätsprinzip (z.B. Sadler et al., 2011) folgend macht innerhalb von Dyaden das Verhalten der einen Person bei der jeweils anderen Person bestimmte Verhaltensweisen besonders wahrscheinlich: sog. komplementäre Verhaltensweisen. Dabei ist auf der Communion-Dimension gleichgerichtetes Verhalten komplementär –z.B. begünstigt warmherziges Verhalten des Lernenden warmherziges Verhalten der Lehrkraft. Im Unterschied dazu ist auf der Agency-Dimension entgegengesetztes Verhalten komplementär - z.B. begünstigt Passivität oder geringe Kompetenz des Lernenden Initiative und Lenkung der Lehrkraft.

Wir haben die Annahme geprüft, dass deshalb Schüler:innen, die noch nicht gelernt haben, sich sozial kompetent gegenüber der Lehrkraft zu verhalten, also z.B. sich nicht kooperativ oder respektvoll zeigen, ein erhöhtes Risiko haben, bei ihrer Lehrkraft wenig kommunales (komplementäres) Verhalten und – unabhängig von ihren fachlichen oder lernmethodischen Kompetenzen - ein stark agentisches (komplementäres) Verhalten sowie negative Emotionen auszulösen.

In einer Beobachtungsstudie schätzten 34 Beobachtende im Klassenzimmer das Verhalten jeweils einer Lehrkraft gegenüber jeweils zwei Schüler:innen der Klasse sowie das Sozialverhalten der Schüler:innen gegenüber der Lehrkraft (Beispielitems: "X verhält sich gegenüber der Lehrkraft kooperativ", "X verhält sich ... feindselig") und – als Maß für ihre Kompetenz – die lernmethodischen

Kompetenzen der Schüler:innen (Beispielitems: "X arbeitet selbständig"; "X kann sich konzentriert mit einer Sache beschäftigen") ein (68 Lehrkraft-Schüler:in-Dyaden). Erwartungsgemäß war das Verhalten der Lehrkraft umso weniger kommunal, je weniger sozial kompetent sich der Lernende ihr gegenüber verhielt, und zwar auch dann, wenn für die lernmethodischen Kompetenzen des Lernenden kontrolliert wurde.

In einem Experiment wurde 102 Lehramtsstudierenden die Vignette eines fiktiven Schülers vorgelegt, der über moderat stark ausgeprägte lernmethodische Kompetenzen (als Maß für seine Kompetenz) beschrieben wurde. Zwischen Versuchspersonen wurde variiert, ob das Sozialverhalten des Schülers gegenüber der Lehrkraft eher ungünstig oder aber unauffällig war. Die Studierenden sollten auf 16 Items, die Verhaltensweisen mit unterschiedlich starker Communion und Agency beschrieben, einschätzen, wie sie sich selbst gegenüber dem Schüler verhalten würden. Erwartungsgemäß beschrieben sie ihr Verhalten gegenüber dem Schüler mit ungünstigem Sozialverhalten als weniger kommunal als gegenüber dem Schüler mit unauffälligem Sozialverhalten. Weiter antizipierten sie erwartungsgemäß, sich gegenüber dem Schüler mit ungünstigem Sozialverhalten stärker agentisch zu verhalten, obwohl sich die Schüler beider Bedingungen in ihren Kompetenzen (von der die Agency der Lehrkraft in einer qualitätsvollen Beziehung ausschließlich abhängig sein sollte) nicht voneinander unterschieden.

In einem weiteren Experiment beschrieben 73 Lehramtsstudierende ihre Emotionen und Zufriedenheit gegenüber demselben fiktiven Schüler, der mit einer moderat stark ausgeprägten lernmethodischen Kompetenz und – experimentell variiert – über eher ungünstiges oder unauffälliges Verhalten gegenüber der Lehrkraft beschrieben wurde. Erwartungsgemäß berichteten die Lehramtsstudierenden gegenüber dem Schüler mit ungünstigem Sozialverhalten weniger positive Emotionen und eine geringere Zufriedenheit in der Beziehung als gegenüber dem Schüler mit unauffälligem Sozialverhalten.

Unsere Befunde sprechen dafür, dass sozial ungünstiges Verhalten gegenüber der Lehrkraft einen Risikofaktor für die schulische Entwicklung des Lernenden darstellt. Wir diskutieren, ob Verhalten und Emotionen von Lehrkräften gegenüber sozial noch wenig kompetenten Schüler:innen verändert werden können, indem sie über diese Zusammenhänge aufgeklärt werden.

E-10: Überzeugungen von Lehrkräften

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S05 R03 H20
Chair der Sitzung: Susanne Geyer

„Mir ist in meinem Anfangsunterricht wichtig, dass ich den Kindern auf keinen Fall den Spaß am Schreiben(lernen) verderbe.“ - Überzeugungen von Grundschullehrkräften zum Schriftspracherwerb

Susanne Geyer, Andreas Hartinger
Universität Augsburg, Germany

Grundschullehrkräften kommt beim Lesen und Schreiben lernen eine Schlüsselrolle zu. Angenommen wird, dass die Art und Qualität von schriftsprachlichen Lernangeboten im Anfangsunterricht maßgeblich durch ihre professionelle Kompetenz bestimmt wird (Vieluf et al. 2020). Neben professionellem Wissen und Können gelten zunehmend auch die Überzeugungen, die ihrem Unterrichtshandeln zugrunde liegen, als wichtige Bestandteile ihrer Expertise (z.B. Schmidt & Schindler 2020; Sturm, Lindauer & Sommer 2019). Allgemein können Überzeugungen als kognitive Filter verstanden werden, durch welche Erfahrungen bewertet und strukturiert werden (Reusser & Pauli 2014). Mehrere Studien konnten in unterschiedlichen Domänen zeigen, dass Überzeugungen maßgeblichen Einfluss u.a. auf die Unterrichtsgestaltung, den Lernerfolg und die Motivation von Schüler*innen haben, moderiert über ihren Einfluss auf Wahrnehmung und Verhalten im Unterricht (vgl. Kunter et al. 2020). Im Kontext des Schriftspracherwerbs wurden diese Zusammenhänge für Grundschullehrkräfte bislang nicht untersucht. Deshalb soll sich unsere Studie dieser Forschungslücken annehmen. Unsere Forschungsidee richtet sich auf das Aufdecken möglicher Zusammenhänge zwischen den Überzeugungen von Grundschul-lehrkräften zum Schriftspracherwerb und der Leistungs- bzw. Persönlichkeitsentwicklung der Schüler*innen am Ende der ersten Jahrgangsstufe. Eine erste Teilfragestellung, die sich in diesem Zusammenhang ergibt, ist, ein geeignetes quantitatives Messinstrument zur Befragung von Lehrkräften hinsichtlich ihrer Überzeugungen zum Schriftspracherwerb zu entwickeln. Fokussiert wird dabei auf das Schreiben – hier verstanden als „eine spezifische sprachliche Handlung, die im Kern auf die selbstständige Produktion von kommunikativ angemessenen und inhaltlich bedeutungsvollen Texten in konkreten Anwendungssituationen abzielt“ (Bachmann & Becker-Mrotzek 2017, S. 25).

Da das Schreiben verschiedene Facetten hat, ist es sinnvoll, zunächst zwischen hierarchieniedrigen Teilprozessen zu unterscheiden (vgl. Sturm, Nänny & Wyss 2017). Hierzu zählen die automatisierbaren Teilkomponenten 1) handschriftliches Schreiben und 2) orthographisch richtiges Schreiben, sowie auch mental anspruchsvollere Aspekte wie 3) eigenes, flüssiges Formulieren. Letzteres wiederum wird im Anfangsunterricht durch das Konzept des „Freien Schreibens“ unterstützt, welches die frühe Einsicht in die kommunikative Funktion von Schreiben und die hohe Relevanz frei von Kindern verfasster, individuell bedeutsamer Texte zugunsten der Schreibmotivation betont (Leßmann 2020; Brügelmann & Brinkmann 2012). Bislang weisen qualitative Befunde aus dem deutschsprachigen Raum darauf hin, dass Grundschullehrkräfte über sehr unterschiedliche Überzeugungen zum Lehren und Lernen im Kontext Orthographie verfügen sowie eine große Unsicherheit im Umgang mit Schreibfehlern spüren (Tressel 2018; Sauerborn 2016). Auch eine Befragung von 44 Primarstufenlehrpersonen aus dritten und vierten Jahrgangsstufen (Wiprächtiger-Geppert et al. 2022) zeigt ein heterogenes Bild insbesondere hinsichtlich der zweiphasigen

didaktischen Modellierung des Unterrichts (zuerst lauttreues Schreiben später orthographisch korrektes Schreiben). Da bislang allerdings nur wenige Untersuchungen vorliegen, die sich mit Lehrkräfteüberzeugungen zum Beginn des Schriftspracherwerbs beschäftigen, besteht hier ein dringender Forschungsbedarf.

Im ersten Schritt soll in Anlehnung an gängige Typologisierungen (z.B. von Merk 2020) überprüft werden, ob sich (wie in anderen Fächern) Überzeugungen 1) zur persönlichen und gesellschaftlichen Bedeutsamkeit des Unterrichtsgegenstands (Orthographie, Textproduktion, Handschreiben), 2) zu den Zielsetzungen für den eigenen Unterricht, 3) zum Lehren und Lernen, 4) zum Zeitpunkt des Beginns verschiedener Lernangebote sowie 5) zu Fehlschreibungen bei eigenen Kinderverschriftungen nicht nur konzeptuell, sondern auch empirisch abbilden lassen. Für die Erfassung dieser Aspekte und Dimensionen wurden teilweise bestehende Items (z.B. zum Lehren und Lernen von Kleickmann 2008) für den Schriftspracherwerb adaptiert oder übernommen (Wiprächtiger-Geppert et al. 2022; Poloczek 2011) sowie eigene Items in Anlehnung an qualitative Studien (Tressel 2018; Sauerborn 2016) entwickelt.

Im geplanten Vortrag wird der Aufbau eines Online-Lehrkräftefragebogens vorgestellt und erste Ergebnisse der Pilotierung berichtet. Die Pilotierung des Instruments wurde an einer Stichprobe von 245 Studierenden des Lehramts Grundschule (2. Fachsemester) vorgenommen. Die meisten Items zu den Überzeugungen wurden mittels fünfstufigen Likert-Skalen erfasst (Antwortoptionen „stimmt gar nicht“ bis „stimmt völlig“). Die Unterrichtsziele wurden mittels eines Dominanzpaar-vergleichs erfragt (vgl. Doering & Bortz 2016). Das adaptierte bzw. neu entwickelte Instrument weist insgesamt eine zufriedenstellende psychometrische Qualität auf.

Instruktionale Überzeugungen und Unterrichtswahrnehmung zu Lernunterstützung von Lehramtsstudierenden am Beispiel Mathematikunterricht der Primarstufe kombiniert entwickeln

Philip Hörter[1], Christina Gippert[2], Manfred Holodynski[2]
[1]Freie Universität Berlin, Germany; [2]Westfälische Wilhelms Universität, Germany

In der ersten Phase der Lehrkräftebildung sollen zukünftige Lehrpersonen auf die spätere Berufspraxis vorbereitet und befähigt werden, im eigenen Unterricht adaptive Lernunterstützung realisieren zu können. Zugehöriges Professionswissen setzt sich u.a. aus verschiedenen Domänen und berufsbezogenen Orientierungen zusammen (Baumert et al., 2006; Dohrmann et al., 2012). Zu Letzteren zählen instruktionale Überzeugungen, die Vorstellungen über Lehren und Lernen von bspw. Fachunterricht (Voss et al., 2011) oder in inklusiven Lernumgebungen (Bertram et al., 2020) beinhalten. Es wird angenommen, dass diese Dispositionen auf Seite der Lehrpersonenebene anhand situationsspezifischer Fähigkeiten, wie *interpretieren*, *bewerten* und *entscheiden* in Performanz überführt wird (Blömeke et al., 2015). Solche Fähigkeiten werden auch als Teacher Noticing bzw. Unterrichtwahrnehmung bezeichnet (Kaiser et al., 2015). Anhand der modelltheoretischen Implikationen lässt sich somit nicht nur für Wissen, sondern auch für Überzeugungen eine Filterwirkung bzw. Guide-Funktion (Fives & Buehl, 2012; Philipp, 2007) für die Wahrnehmung annehmen, die damit auf den Unterricht und Lernendenleistung wirken (geht hier Blömeke et al., 2022). Meschede et al. (2017) nehmen darüber hinaus einen bi-direktionalen

Zusammenhang an, nach dem Unterrichtswahrnehmung gleichermaßen auf Wissenserwerb oder Überzeugungsveränderung wirken kann.

Diese Vermutung aufgreifend wurde in der vorgestellten Studie untersucht, inwiefern ein Universitätsseminar zu Lernunterstützung am Beispiel des Mathematikunterrichts der Primarstufe die Fähigkeiten zur Unterrichtswahrnehmung fördert und diese Förderung zugleich mit einer Adressierung instruktionaler Überzeugungen erfolgreich im Sinne einer Veränderung von instruktionalen Überzeugungen einhergeht. Grundlage waren Unterrichtsclips aus dem Videoportal *ProVision*.

Die 31 teilnehmenden Studierenden der Experimentalgruppe zeigten nach dem Trainingsseminar eine signifikante Verbesserung der inhaltsgeleiteten Interpretation von Unterrichtsereignissen im Vergleich zu einer Gruppe mit dem inhaltlichen Fokus Klassenführung (n =36) und einer Vergleichsgruppe ohne spezifisches Training (n = 72). Zudem konnten für die Experimentalgruppe Veränderungen (Prä-Post-Follow-up) in den instruktionalen Überzeugungen anhand von *Conditional Growth Curve* - Modellen festgestellt werden, die aber nicht von den anderen beiden Gruppen signifikant verschieden waren. In der Experimentalgruppe konnte zudem eine Korrelation zwischen einem Zuwachs in der Fähigkeit der Situationsbewertung und einem stärkeren Rückgang instruktional-transmissiver Überzeugungen festgestellt werden.

Zusammengefasst lassen sich Unterrichtsvideos gezielt zur Entwicklung fokusspezifischer Unterrichtswahrnehmungsfähigkeiten einsetzen, die mit – eher trainingsunabhängigen – Veränderungen von instruktionalen Überzeugungen jedoch in leichter Wechselwirkung stehen. Zwar gelten derartige Überzeugungen damit weiterhin als relativ stabil (Philipp, 2007) gegenüber direkter Beeinflussung, sollten jedoch als Teil der Professionsentwicklung expliziert und adressiert werden, wie bei der Analyse und Reflexion von Unterricht anhand von Unterrichtsvideos möglich.

Die Bedeutung von Einstellungen zum bildungswissenschaftlichen Studium für die Entwicklung pädagogisch-psychologischer Kompetenzen von Lehramtsstudierenden

Daniel Mann, Jennifer Paetsch, Jakob Meyer
Otto-Friedrich-Universität Bamberg, Germany

Die professionelle Kompetenz einer Lehrkraft gilt als wichtige Voraussetzung für professionelles Handeln im Unterricht (Baumert & Kunter, 2006; Kunter, Kleickmann, Klusmann & Richter, 2011). Kompetenzunterschiede von Lehrkräften können dabei sowohl auf ihre Qualifizierung als auch auf Unterschiede in individuellen Voraussetzungen zurückgeführt werden.

Dieser Beitrag untersucht, wie sich individuelle Voraussetzungen von Lehramtsstudierenden auf die Entwicklung von selbsteingeschätztem sowie objektiv gemessenem Wissen auswirken. Dabei wird auf die Entwicklung des pädagogisch-psychologischen Wissens als ein Teilaspekt der Lehrkräftekompetenz fokussiert. Dieses Wissen ist fachübergreifend und von allen Lehramtsstudierenden zu erwerben. Die fachübergreifenden Studieninhalte, die sich auf die Themenschwerpunkte Erziehung und Bildung, Schule und Unterricht sowie Lernen und Sozialisation beziehen, werden im Rahmen des bildungswissenschaftlichen Studienanteils vermittelt (Kunina-Habenicht et al., 2012).

Bisherige Untersuchungen zeigen, dass die von den Studierenden wahrgenommenen bildungswissenschaftlichen Lerngelegenheiten erheblich variieren (Hohenstein et al., 2014). Es wird daher angenommen, dass die individuelle Gestaltung und Nutzung der Lerngelegenheiten im bildungswissenschaftlichen Studium den Kompetenzerwerb der Studierenden beeinflussen (Kunina-Habenicht et al. 2013). Dabei wird positiven Einstellungen zu relevanten Merkmalen des fächerübergreifenden Studienanteils (z. B. Datennutzung, Forschung, Praxisbezug) eine wichtige Bedeutung zugesprochen (Thoren, Wißmann, Kinder & Hannover, 2018).

Fragestellung

In dieser Studie soll geprüft werden, wie sich die Einstellungen zum Praxis- und Forschungsbezug des bildungswissenschaftlichen Studienanteils und die subjektiv angegebene Häufigkeit genutzter Lerngelegenheiten auf die Entwicklung der selbsteingeschätzten sowie objektiv erfassten pädagogisch-psychologischen Kompetenzen auswirken.

Methode

Die Daten der vorliegenden Längsschnittstudie wurden während des Wintersemesters 2018/19 (MZP1) und am Ende des Wintersemester 2019/20 (MZP2) erhoben.

Die Stichprobe besteht aus 78 Lehramtsstudierenden, die mindestens an zwei Messzeitpunkten teilgenommen haben. Die Studierenden verteilen sich wie folgt auf die Schularten: Grundschule 63 %, Mittelschule/Realschule 15 %, Gymnasium 13 %, Berufliche Bildung 8 %, und waren im Mittel 20.4 Jahre alt (*SD* = 2.5). Die Mehrheit der Teilnehmenden (89 %) ist weiblich.

Zur Erfassung der selbsteingeschätzten pädagogisch-psychologischen Kompetenzen wurde zu beiden Messzeitpunkten ein Instrument aus 6 Skalen von Thiel und Blüthmann (2009) eingesetzt (Subskalen .71 ≤ α ≤ .91). Weiterhin wurden selbsteingeschätzte didaktische Kompetenzen durch 14 Items (Thiel & Blüthmann, 2009, α = .94) erfasst.

Das objektive bildungswissenschaftliche Wissen wurden durch einen standardisierten Test (BwW-Kompetenztest SPEE, Seifert & Schaper, 2012) gemessen.

Zur Erfassung der Einstellungen der Studierenden zum bildungswissenschaftlichen Studium wurde ein Instrument neu entwickelt. Es erfasst, welche Einstellungen die Studierenden einerseits zum Praxis- (3 Items, α = .78) und andererseits zum Forschungsbezug (3 Items, α = .70) in ihrem bildungswissenschaftlichen Studium haben.

Die von den Studierenden wahrgenommenen Lerngelegenheiten wurden ebenfalls durch ein neu entwickeltes Instrument auf Grundlage von BilWiss (Terhart et al., 2012) sowie der KMK-Standards für die Lehrerbildung mit 70 Items in 8 Themenbereichen erhoben.

Ergebnisse

Die deskriptiven Ergebnisse zu den Einstellungen zum bildungswissenschaftlichen Studium zeigen, dass die Studierenden durchschnittlich hoch ausgeprägte positive Einstellungen zum Praxisbezug aufweisen; die Einstellungen zum Forschungsbezug weisen hingegen weniger starke Zustimmungsraten auf (M_{Prax} = 3.73; M_{Fors} = 2.65 bei 4-stufiger Antwortskala). Ergebnisse aus *t*-Tests zeigen einen signifikanten Zuwachs von MZP1 zu MZP2 für alle Selbsteinschätzungen des

pädagogisch-psychologischen Wissens bzw. didaktischer Kompetenzen sowie das objektive Wissen (SPEE).

Die Fragestellung wurde mit linearen Regressionsanalysen bearbeitet. Als Prädiktoren wurde jeweils das selbsteingeschätzte Wissen zu *t1*, der Summenscore genutzter Lerngelegenheiten, sowie die Einstellungen zum Praxisbezug und Forschungsbezug aufgenommen. Die Ergebnisse zeigen, dass die Einstellung zum Forschungsbezug signifikant Varianz im Wissenszuwachs für die Skalen „individuelle Förderung“ und „Leistungsbeurteilung“ aufklärt. Für die Skala „Lerndiagnostik“ wird das Signifikanzniveau knapp verfehlt. Die Einstellung zum Praxisbezug ist ein signifikanter Prädiktor für die Skala „Klassenführung“. Die Summe der genutzten Lerngelegenheiten leistet keinen signifikanten Beitrag in der Varianzaufklärung. Weiterhin ließen sich keine signifikanten Prädiktoren für das objektive Wissen nachweisen (SPEE).

In noch ausstehenden Analysen sollen weitere individuelle Merkmale berücksichtigt werden (Schulform, Praxiserfahrung).

Motivationsprofile zum Erwerb pädagogischen Wissens im Lehramtsstudium

Martina Homt, Janis Wehde, Bea Bloh
Universität Paderborn, Deutschland

Für das professionelle Handeln von Lehrkräften wird eine evidenzbasierte Praxis, also die situationsangemessene Berücksichtigung von wissenschaftlichen Befunden und bewährten Theorien, gefordert (Bauer, Prenzel & Renkl, 2015). Entsprechend betont auch die Kultusministerkonferenz (KMK, 2019) die Bedeutung wissenschaftlichen Wissens in der Lehrer*innenbildung für die zukünftige Aufgabenwahrnehmung. Für angehende Lehrkräfte wird neben dem Fachwissen und dem fachdidaktischen Wissen das pädagogische Wissen als wichtiger Bestandteil des Professionswissens gefasst (z.B. Baumert & Kunter, 2006; Shulman, 1987), wenngleich der Begriff pädagogisches Wissen nicht übereinstimmend verwendet wird. Voss, Kunina-Habenicht, Hoehne und Kunter (2015) systematisieren die existierende Literatur und Forschungsaktivitäten im deutschsprachigen Raum und definieren pädagogisches Wissen als „Kenntnisse über das Lernen und Lehren, die sich auf die Gestaltung von Unterrichtssituationen beziehen und die fachunabhängig, das heißt auf verschiedene Fächer und Bildungsbereiche anzuwenden sind“ (S. 194).

Werden die Befunde zum Erwerb pädagogischen Wissen von (angehenden) Lehrkräften betrachtet, zeigt sich übereinstimmend, dass dieses im Studienverlauf zunimmt (z.B. Blömeke, Müller, Felbrich & Kaiser, 2008; König & Klemenz, 2015). Jedoch ist bisher wenig über den Prozess dieses Wissenserwerbs im Studium bekannt und insbesondere „welche Elemente der Lehrerbildung besonders hilfreich sind“ (Voss et al., 2015, S. 208). Studierende profitieren sehr unterschiedlich von Lerngelegenheiten und die Annahme ist, dass je nach affektiv-motivationalen Voraussetzungen die Lerngelegenheiten zum Erwerb von pädagogischem Wissen unterschiedlich effektiv genutzt werden (ebd.). Was fehlt, sind spezifischere Untersuchungen, die affektiv-motivationale Faktoren mit Blick auf Prozesse des pädagogischen Wissenserwerbs im Lehramt fokussieren. Daher soll im Beitrag der Frage nachgegangen werden, welche verschiedenen Motivationsgruppen sich für den Erwerb pädagogisch-psychologischen Wissens identifizieren lassen.

Um die affektiv-motivationalen Voraussetzungen der Studierenden für den Erwerb pädagogischen Wissens differenziert abbilden und untersuchen zu können, wurde in Anlehnung an die Theorie der organismischen Integration (Ryan & Deci, 2019) ein Instrument zur Erfassung intrinsischer und extrinsischer Motivationsregulationsformen eingesetzt, das vier Motivationsformen (externale Regulation, introjizierte Regulation, identifizierte Regulation, intrinsische Regulation) als Subskalen konzipiert, die sich in ihrem Grad der Fremd- bzw. Selbstbestimmung unterscheiden. Das Instrument beinhaltet insgesamt 18 Items zur Erfassung der Motivation zum Erwerb pädagogischen Wissens (z.B. Ich beschäftige mich mit Theorien, weil es mir wichtig ist, diese zu lernen.), die jeweils mit einer fünfstufigen Likert-Skala (stimme überhaupt nicht zu bis stimme voll zu) eingeschätzt werden konnten. Es ist in einer vierfaktoriellen Modellierung geeignet, um die Motivationsregulationsformen abzubilden (Grosser & Wehde, 2021).

Auf Basis der Motivationsformen wurden mittels Latent Class Analysis (LCA) in R verschiedene Motivationsprofile Studierender identifiziert und anschließend mit demographischen (z.B. Geschlecht, Migrationshintergrund, Erwerbstätigkeit) und professionsbezogenen Merkmalen (Selbstwirksamkeitserwartung, Gewissenhaftigkeit, Intention zur evidenzbasierten Praxis) in Beziehung gesetzt. Diese Gruppenvergleiche erfolgten mittels varianzanalytischer Verfahren.

Die Untersuchung erfolgte an einer Stichprobe bestehend aus N = 350 Lehramtsstudierenden der Sekundarstufe I (73,7 % weiblich, 25,7 % männlich, 0,6 % keine Angabe). Das durchschnittliche Alter der befragten Lehramtsstudierenden beträgt 25,94 Jahre (SD = 3,27). Die Datenerhebung erfolgte computergestützt und anonymisiert in fünf Kohorten (Sommersemester 2020 bis Sommersemester 2022) jeweils in der ersten Sitzung einer Lehrveranstaltung zum pädagogischen Wissen im Master of Education an der Universität Paderborn.

Die LCA ergab die besten Modellfits (BIC, AIC) für die Fünf-Gruppenlösung. Die Gruppengrößen variieren von 47 bis 116. Neben den zwei Gruppen, die im Profil klassisch ein intrinsisches bzw. extrinsisches Profil aufweisen, zeigen sich in den drei anderen Gruppen stärkere Mischformen. Die Gruppen unterscheiden sich hinsichtlich demographischer, aber auch professioneller Merkmale. Zudem zeigt sich, dass die identifizierte Motivation im Durchschnitt die stärkste Motivationsform ist. Die Ergebnisse werden im Hinblick auf Strukturmerkmale universitärer Lehrerausbildung, didaktische Arrangements und einen lebenslangen Professionalisierungsprozesses diskutiert.

Symposien

E-11: Lernen auf Abruf? Balancieren zwischen Anstrengung und Abruferfolg

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* S07 S00 D07

Lernen auf Abruf? Balancieren zwischen Anstrengung und Abruferfolg

Chair(s): **Roman Abel** (Ruhr-Universität Bochum), **Julian Roelle** (Ruhr-Universität Bochum)
*Diskutant*in(nen):* **Detlev Leutner** (Universität Duisburg-Essen)

Die nachhaltigen Lernvorteile, die auf einen Abruf von Wissen aus dem Gedächtnis zurückzuführen sind, gelten in der Bildungsforschung als unbestritten. Die Frage nach der optimalen Gestaltung von Abrufübungsaufgaben im Bildungskontext ist jedoch keineswegs geklärt. Bisherige Befunde deuten darauf hin, dass die Effektivität von Abrufübung substanziell davon abhängt, wieviel Anstrengung die Lernenden investieren müssen, um die jeweiligen Gedächtnisinhalte abzurufen und wie erfolgreich sie dabei sind. Dabei sind die beiden Gelingensbedingungen gegenläufig: Je mehr Anstrengung es kostet, desto unwahrscheinlicher ist es, die Gedächtnisinhalte korrekt und vollständig abzurufen. Die vier Beiträge des Symposiums beziehen sich allesamt auf die Untersuchung dieser Gelingensbedingungen, unterscheiden sich jedoch darin, an welcher der beiden Stellschrauben die zu untersuchenden Lösungsansätze ansetzen.

Beitrag 1 zielt darauf ab, vor der Abrufübung Wissenskonstruktionsprozesse mittels Beispielgenerierungsaufgaben anzuregen, um den nachfolgenden Erfolg von Abrufübungen zu steigern. Es wird betrachtet, ob das Generieren von Beispielen vor der Abrufübung den Erfolg der Abrufübung sowie den Lernerfolg im Vergleich zum Generieren von Beispielen nach der Abrufübung und im Vergleich zu Abrufübung ohne Beispielgenerierung steigert.

Beiträge 2 bis 4 zielen darauf ab, Wissenskonstruktionsprozesse nicht nur vor, sondern während der Abrufphase anzuregen, um die investierte Anstrengung zu steigern. In Beitrag 2 werden Wissenskonstruktionsprozesse angeregt, indem die Lernenden selbständig Fragen zu den Lerninhalten generieren. Dabei wird manipuliert, ob das Lernmaterial während des Generierens der Fragen und während des Beantwortens jeweils verfügbar ist oder nicht verfügbar ist (*Abrufübung*).

In Beitrag 3 werden Wissenskonstruktionsprozesse beim Abruf mittels des Verschachtelns (*Interleaving*) angeregt. Es wird untersucht, ob ein verschachtelter Abruf von Lerninhalten, die in einem zuvor gelesenen Lerntext geblockt präsentiert wurden, den Lernerfolg gegenüber einem geblockten Abruf steigert.

In Beitrag 4 werden Wissenskonstruktionsprozesse durch Abrufaufgaben angeregt, die es erforderlich machen, multiple Gedächtnisinhalte zu verknüpfen. Es wird untersucht, ob Abrufaufgaben, die es erfordern, Gedächtnisinhalte miteinander zu verknüpfen, solchen Aufgaben, die den Abruf isolierter Gedächtnisinhalte erfordern, überlegen sind.

Das vorliegende Symposium zeichnet sich durch die Diversität zu untersuchender Lösungsansätze aus, Anstrengung und Abruferfolg in Balance zu bringen. Bei allen Beiträgen handelt es sich um experimentelle Studien, die expositorische Lerntexte nutzen und den Lernerfolg nach einer Woche Verzögerung erfassen. Der Diskutant Prof. Dr. Dr. h.c. Detlev Leutner ist ein international herausragender Experte im Bereich des Lernens durch Wissenskonstruktion. Seine Sichtweise auf die vorgestellten Forschungsarbeiten wird daher sicherlich dazu beitragen, fruchtbare

Schlussfolgerungen im Hinblick auf den Nutzen von Wissenskonstruktionsaktivitäten für die Gestaltung effektiver Abrufübungen zu ziehen.

Beiträge des Symposiums

Kann eine sequenzielle Kombination von Wissenskonstruktions- und Abrufübungslernaufgaben den Erwerb deklarativer Konzepte gewinnbringend fördern?

Niklas Obergassel[1], Alexander Renkl[2], Tino Endres[2], Matthias Nückles[2], Shana Carpenter[3], Julian Roelle[1]
[1]Ruhr-Universität Bochum, [2]Albert-Ludwigs-Universität Freiburg, [3]Iowa State University

Theoretischer Hintergrund

Eine vielversprechende instruktionale Maßnahme zur Unterstützung des Wissenserwerbs liegt in der Darbietung von Wissenskonstruktionslernaufgaben (Aufgaben, welche die aktive Ausführung von Wissenskonstruktionsprozessen anregen; siehe Brod, 2021) oder Abrufübungslernaufgaben (Aufgaben, welche den aktiven Informationsabruf aus dem Gedächtnis anregen; siehe McDermott, 2021). So zeigt die Empirie, dass beide Aufgabenarten den Erwerb von Wissen über Inhalte, die in einer vorherigen Lesephase eingeführt wurden, bedeutend fördern können (Agarwal et al., 2021; Rawson & Dunlosky, 2016). Diese Gemeinsamkeit in Punkto Effektivität soll jedoch keinesfalls darüber hinwegtäuschen, dass sich beide Aufgabenarten aus theoretischer Sicht funktionell stark unterscheiden: So zielen Wissenskonstruktionslernaufgaben primär auf die Bildung kohärenter mentaler Repräsentationen zur Verständnisförderung ab (Fiorella & Mayer, 2016), während Abrufübungslernaufgaben hauptsächlich der Konsolidierung bestehender mentaler Repräsentationen und damit einer verbesserten Behaltensleistung Rechnung tragen (Karpicke & Grimaldi, 2012). Ausgehend von den skizzierten Funktionen und der Annahme, dass sich diese perfekt ergänzen sollten, drängt sich der Gedanke auf, dass eine (sequenzielle) Kombination beider Aufgabenarten den Wissenserwerb noch effektiver fördern könnte als jede Aufgabenart alleine. Überraschenderweise hat solch eine Kombination jedoch bisher wenig Beachtung gefunden. So haben zwar einige, wenige Studien (O'Day & Karpicke, 2021; Roelle et al., 2022) erste Anstrengungen unternommen, die Effekte einer solchen Kombination zu explorieren, jedoch aufgrund methodischer Einschränkungen mehr Fragen als Antworten geliefert.

Fragestellung

Vor diesem Hintergrund war es das Ziel der Studie, am Beispiel des Lernens deklarativer Konzepte (Lernen von abstrakten Konzepten und ihrer Definitionen) einen Beitrag zur Forschung im Bereich der (sequenziellen) Kombination von Wissenskonstruktions- und Abrufübungslernaufgaben zu leisten. Konkret wurde untersucht, ob (a) eine sequenzielle Kombination beider Aufgabenarten (Wissenskonstruktion-zuerst- oder Abrufübung-zuerst-Sequenz) mit einem höheren Lernerfolg verbunden sein würde als Wissenskonstruktionslernaufgaben (FF 1) oder Abrufübungslernaufgaben alleine (FF 2) und (b), ob sich die zwei kombinatorischen Sequenzen bezüglich des Lernerfolgs unterscheiden würden (FF 3).

Methode

Um diese Forschungsfragen zu adressieren, wurde ein Laborexperiment mit Studierenden durchgeführt. Von den 320 Studierenden, die basierend auf einer A-priori-Teststärkenanalyse

(durchgeführt mittels G*Power) als notwendige Stichprobe ermittelt wurden, konnten bisher 108 Studierende (80 weiblich, MAlter = 23.12) rekrutiert werden. Alle Teilnehmenden absolvierten zunächst einen Vortest und lasen einen Text, in dem mehrere deklarative Konzepte eingeführt wurden. Anschließend wurden alle Teilnehmenden randomisiert einer von zehn Bedingungen zugewiesen, die sich aus einem 5x2-Zwischensubjekt-Design mit den Faktoren Lernaufgabensequenz und Nachtestzeitpunkt ergaben. In zwei aufeinanderfolgenden Lernphasen bearbeiteten die Lernenden dann, abhängig vom Faktor Lernaufgabensequenz, entweder (a) nur Wissenskonstruktionslernaufgaben (Nur-Wissenskonstruktion-Sequenz), (b) nur Abrufübungslernaufgaben (Nur-Abrufübung-Sequenz), (c) nur Restudy-Aufgaben (Nur-Restudy-Sequenz; Kontrollsequenz), (d) zuerst Wissenskonstruktionslernaufgaben und dann Abrufübungslernaufgaben (Wissenskonstruktion-zuerst-Sequenz) oder (e) zuerst Abrufübungslernaufgaben und dann Wissenskonstruktionslernaufgaben (Abrufübung-zuerst-Sequenz). Nach Abschluss der Lernphasen bearbeiteten die Lernenden abhängig vom Faktor Nachtestzeitpunkt entweder einen direkten oder einen einwöchig-verzögerten Nachtest.

Vorläufige Ergebnisse und ihre Bedeutung

Im Hinblick auf den Lernerfolg im direkten Nachtest ergaben erste Auswertungen der verfügbaren Daten mittels Kontrastanalyse eine signifikante Überlegenheit der Abrufübung-zuerst-Sequenz gegenüber der Nur-Restudy-Sequenz, $t(17.35) = 3.06$, adjustiertes $p = .032$, $d = 1.23$, der Nur-Abrufübung-Sequenz, $t(17.93) = 2.79$, adjustiertes $p = .042$, $d = 1.16$, sowie der Wissenskonstruktion-zuerst-Sequenz, $t(16.81) = -3.33$, adjustiertes $p = .032$, $d = -1.48$. In Bezug auf den Lernerfolg im verzögerten Nachtest zeigten sich hingegen keine Unterschiede zwischen den Sequenzen, alle adjustierten $p > .05$, $-0.32 \leq d \leq 0.85$.

Die vorläufigen Ergebnisse deuten darauf hin, dass eine Kombination von Wissenskonstruktionslernaufgaben und Abrufübungslernaufgaben durchaus einen Mehrwert für den Lernerfolg bieten kann. Dieser Mehrwert scheint jedoch möglicherweise auf den kurzfristigen Lernerfolg sowie den Vergleich mit Abrufübungslernaufgaben alleine beschränkt zu sein und nur dann aufzutreten, wenn eine Abrufübung-zuerst-Sequenz, die einer Wissenskonstruktion-zuerst-Sequenz im Hinblick auf kurzfristige Effektivität überlegen zu sein scheint, zur Kombination verwendet wird. Eine Analyse der vollständigen Daten, die für tiefergehende Schlussfolgerungen notwendig ist, wird bis zur Konferenz erfolgt sein.

Fragengenerieren als Abrufübung – Wechsel zwischen Generieren und abruforientiertem Beantworten verbessert das Lernen

Tino Endres, Alexander Renkl
Albert-Ludwigs-Universität Freiburg

Die Lernstrategie des Fragengenerierens gehört seit vielen Jahren zum Repertoire von Lernenden. Da diese generative Lernaktivität Fragen beinhaltet, liegt die Kombination mit Abrufübungen, die oft als Testungseffekt bezeichnet werden (Rowland, 2014) nahe. Die Kombination dieser beiden effektiven Strategien ist vielversprechend, um sowohl das Verständnis als auch das Behalten über längere Zeiträume zu verbessern. Leider haben neuere Studien nicht die erwarteten Unterschiede zwischen der Generierung von Testfragen mit und ohne Abrufanforderungen identifizieren können

(z.B. Hoogerheide et al., 2019). Dies könnte daran liegen, dass die Fragengenerierung und die Beantwortung dieser Fragen unterschiedliche Bedingungen benötigen, um ihr volles Potenzial zu entfalten, und daher separat betrachtet werden müssen.

Fragengenerierung

Bei der Generierung von Fragen sollte die Verfügbarkeit des Lernmaterials sowohl die Qualität der Fragen als auch die Abdeckung der Lerninhalte erhöhen. Daher erwarten wir, dass die Lernmaterialverfügbarkeit bei der Fragengenerierung die Qualität der Fragen erhöht. Diese Steigerung der Fragenqualität sollte zu einem höheren Lernerfolg in einem unmittelbaren Lerntest führen.

Beantwortung der generierten Fragen

Abrufeffekte werden oft erst nach einer Verzögerung beobachtbar. Die in das Abrufen investierte Anstrengung erhöht die Behaltensleistung meist nur substanziell bei einem um mindestens eine Woche verzögerten Lerntest (Rowland, 2014). Daher gehen wir davon aus, dass abruforientiertes Beantworten der Testfragen die mentale Anstrengung erhöht und somit die verzögerte Lernleistung steigert.

Besonders interessant ist die Kombination dieser beiden Bedingungen. Bei Abrufübungen ist ein erfolgreicher Abruf essentiell, um Abrufeffekte zu finden. Eine Beschäftigung mit einer konstruktiven Lernaufgabe wie das Fragengerieren vor dem Abruftraining könnte den Abruferfolg erhöhen und einen Testungseffekt sicherstellen. Ein Wechsel zwischen nicht abruforientiertem generieren und abruforientiertem Beantworten könnte beide Prozesse verbessern. Eine solche Abfolge würde eine optimale Leistung bei der Fragenerstellung (mit verfügbarem Lernmaterial) ermöglichen und dennoch der Beantwortung dieser Fragen einen Abruf erfordern. In dieser Studie konzentrierten wir uns auf diese sequentielle Implementierung dieser beiden Aufgaben, um die besten Bedingungen sowohl für generative Prozesse als auch für die Abrufübung zu schaffen.

Methode

Wir untersuchten diese Fragestellung in einem 3x2x2-mixed Design. Im Between-Faktor (3) variierten wir, ob die Lernenden Fragen mit verfügbarem Lernmaterial, ohne das Lernmaterial oder in einer sequentiellen Bedingung generierten und beantworteten. Außerdem variierten wir zwei innersubjekt Faktoren. Erstens wurde die Komplexität der Textabschnitte (2) variiert: Innerhalb eines Lerntextes gab es leichte und komplexe Textpassagen. Zweitens wurde der Zeitpunkt der Bewertung (2) variiert: Für diese Variation verwendeten wir zwei Texte und führten bei einem Text einen unmittelbaren Lerntest und beim anderen Text einen um eine Woche verzögerten Lerntest durch. Wir testeten 85 Studierende online (Mage=27,30, SD=11,77). 78 nahmen an dem verzögerten Posttest teil.

Ergebnisse

Es gab eine signifikante Interaktion zwischen den Bedingungen zwischen den Versuchspersonen und dem Zeitpunkt des Lerntests, $F(2,75)=4,260$, $p=.018$, $\eta 2p=.111$. Wir berechneten Kontraste, um unsere Hypothese zu testen. Die Verfügbarkeit des Lernmaterials erhöhte die Qualität der generierten Fragen signifikant ($p<.05$). sowie die unmittelbare Lerntest ($p<.05$). Die Bedingung mit durchgehender

Verfügbarkeit des Lernmaterials und die sequenzierte Bedingung übertrafen die Bedingung ohne Lernmaterial in diesen Variablen. Eine Mediationsanalyse zeigte einen signifikanten indirekten Effekt über die Qualität der Fragen auf den unmittelbaren Lerntest. Im verzögerten Lerntest schnitt die sequenzierte Bedingung besser ab als beiden anderen Bedingungen.

Diskussion

In dieser Untersuchung führte die sequenzielle Umsetzung der Fragengenerierung zu den besten Lernergebnissen, insbesondere im verzögerten Lerntest. Die Bedingung ohne Lernmaterial während beider Phasen scheint in der anfänglichen Lernphase zu wenig gelernt zu haben, um ihr Potenzial im verzögerten Lerntest entfalten zu können. Nach der Durchführung der generativen Lernaufgabe mit dem vorhandenen Lernmaterial profitierten die Lernenden jedoch von den zusätzlichen Abrufanforderungen. Diese Ergebnisse unterstreichen die Bedeutung des Abruferfolgs bei der Durchführung von Abrufübungen und zeigen, dass die Kombination von generativen Lernaufgaben, wie z.B. Fragengenerierung, den wichtigen Faktor des Abruferfolgs erhöhen kann.

Effekte verschachtelten Abrufs auf Inferenzbildung beim Lesen expositorischer Texte

Roman Abel, Julian Roelle
Ruhr-Universität Bochum

Theoretischer Hintergrund

Aktives Abrufen von Gedächtnisinhalten ist ein effektives Mittel, um nachhaltiges Lernen zu fördern. Der Nutzen einer Abrufübung steigt mit dem Ausmaß an Anstrengung, die in den Abruf investiert wird, sofern der Abruferfolg konstant bleibt (Karpicke, 2017). Ein vielversprechender Ansatz zur Steigerung der Effektivität von Abrufübungen besteht darin, Bedingungen zu schaffen, die das Abrufen anstrengend machen, ohne den Abruferfolg zu verringern.

Über zwei Experimente hinweg kombinieren wir zwei wünschenswerte Erschwernisse, Abrufübung und Verschachteln (Interleaving) (cf. Bjork & Bjork, 2014). Mittels verschachtelter Sequenzen versuchen wir einerseits die Kohärenzbildung vor der Abrufübung zu fördern (um damit den Abruferfolg zu steigern) und andererseits die Anstrengung beim Abruf zu steigern (ohne Einbußen für den Abruferfolg zu verursachen). Dabei untersuchen wir die Auswirkungen dieser Kombinationen auf Wiedergabeleistung und Inferenzbildung.

Fragestellungen

1). Bereitet das Lesen eines Lerntextes, in dem die Informationen verschachtelt präsentiert werden, die Lernenden besser auf den nachfolgenden Abruf vor (Abruf- und Lernerfolg) als das Lesen eines geblockten Textes?

2). Ist eine verschachtelte Abrufsequenz (Lernphase: A1A2A3, B1B2B3, C1C2C3; Abrufphase: A1B1C1, A2B2C2, A3B3C3) vorteilhafter als eine blockierte Abrufsequenz (d.h. Lern- und Abrufphase: A1A2A3, B1B2B3, C1C2C3)?

Methode

Wir führten parallel zwei between-subjects Experimente mit Studierenden durch. Von 180 Studierenden für Experiment 1 und 160 Studierenden für Experiment 2, die als notwendige

Stichprobengrößen bestimmt wurden, konnten bisher 95 Studierende für Experiment 1 und 96 Studierende für Experiment 2 getestet werden.

In Experiment 1 lasen die Lernenden einen geblockten Text über sechs Wale (d.h. sechs Merkmale von Wal 1, sechs Merkmale von Wal 2, ...) und riefen den Inhalt in geblockter Sequenz ab (zuerst alle sechs Merkmale von Wal 1, dann alle sechs Merkmale von Wal 2, ...) vs. riefen den Inhalt in verschachtelter Sequenz ab (d.h. zuerst Merkmal 1 von allen sechs Walen, dann Merkmal 2 für alle sechs Wale, ...) vs. lasen den Text erneut (Kontrollgruppe). In Experiment 2 manipulierten wir die Textsequenz (geblockt vs. verschachtelt) und die Abrufsequenz (geblockt vs. verschachtelt). Eine Woche später wurden Wiedergabe und Inferenzbildung gemessen (darunter Vergleiche zwischen Walen und Erkennen von Gesetzmäßigkeiten, die im Text nicht berichtet wurden wie z.B., dass nur Bartenwale zwischen Äquator und Polarregion wandern, aber Zahnwale nicht). Die Textsequenzmanipulation sowie Inferenzmessung können Abel et al. (2020) entnommen werden.

Vorläufige Ergebnisse

In Experiment 1 förderte der verschachtelte Abruf die Wiedergabeleistung (relativ zum geblockten Abruf, $p = .027$) und das Erkennen von Gesetzmäßigkeiten ($p = .047$). Im Einklang mit diesen Ergebnissen führte der verschachtelte Abruf (relativ zum geblockten Abruf) in Experiment 2 dazu, dass die Gesetzmäßigkeiten besser erkannt wurden ($p = .036$).

Diskussion

Ausgehend von vorläufigen Ergebnissen kann man festhalten, dass über beide Experimente hinweg der verschachtelte Abruf das Erkennen von Gesetzmäßigkeiten und, zumindest in Experiment 1, die Wiedergabeleistung im Vergleich zum geblockten Abruf förderte. Bisher konnten wir keinen Nachweis für die Wirksamkeit von verschachtelten Texten zur besseren Vorbereitung auf die anschließende Abrufübung entdecken. Die Ergebnisse deuten darauf hin, dass Verschachtelung des Abrufs eine vielversprechende Möglichkeit darstellt, den Nutzen von Abrufübungen zu steigern.

The best of both worlds: Abrufaufgaben mit Fokus auf Verständnis fördern Wiedergabe und Verständnis

Sophia C. Weissgerber[1], Roman Abel[2]
[1]Universität Kassel, [2]Ruhr-Universität Bochum

Theoretischer Hintergrund

Um maximalen Praxisnutzen aus Abrufübungen zu ziehen, sollte man deren Wirkmechanismen verstehen. Theorien zur Erklärung des Testungseffektes wie effortful retrieval hypothesis (Carpenter & Losch, 2006; Endres et al., 2017) führen den Lernvorteil des aktiven Abrufs aus dem Gedächtnis auf die Anstrengung zurück, die investiert wird, um die entsprechenden Gedächtnisinhalte zu identifizieren (effortful memory search), was zur Folge hat, dass die abgesuchten Gedächtnisinhalte aktiviert und miteinander vernetzt werden (Chan et al., 2006). Solche Ansätze wie transfer appropriate processing account (McDaniel & Butler, 2011) und relevance instruction account (McCrudden & Schraw, 2007) betonen dagegen die Bedeutung der Passung zwischen Wissenskonstruktionsprozessen, die einerseits in der Lern- bzw. der Abrufphase und andererseits im Lerntest verlangt werden. Wenn sich die Wissenskonstruktionsprozesse überlappen, sollte Abruf lernförderlich sein. Ausgehend von der Diskrepanz der Erklärungsansätze, ist es für die Praxis

unersichtlich, ob der Abruf notwendig ist, um Wiedergabe und oder Verständnis zu fördern, und wie ggf. eine Abrufübung gestaltet werden sollte, um beides zu fördern.

Fragestellungen

Die vorliegende Studie untersucht, ob Aufgaben, die auf das Verknüpfen multipler Textinhalte abzielen (Fokus auf Verständnis von Zusammenhängen) solchen Aufgaben überlegen sind, die auf die Wiedergabe isolierter Textinhalte abzielen (Fokus auf Wiedergabe). Insbesondere wird geschaut, ob solche Abrufaufgaben, die Schlussfolgerungen erfordern, solchen Abrufaufgaben überlegen sind, die lediglich den Abruf isolierter Gedächtnisinhalte erfordern.

Methode

Zur Untersuchung dieser Fragestellung haben wir ein Experiment mit Studierenden durchgeführt (N = 190), welche zunächst einen anspruchsvollen expositorischen Text (bestehend aus 15 Abschnitten) zum physikalischen Phänomen der Interferenz lasen. Beim zweiten Lesen kamen im Anschluss an jeden von insgesamt 15 Textabschnitte Aufgaben zum Einsatz. Die Gestaltung der Aufgaben variierte zwischen den Versuchspersonen in Bezug auf a) den Vernetzungsgrad von Textinformation, welche durch eine Aufgabe abgedeckt wurden (isolierte Fakten vs. Zusammenhänge) und b) ob die Aufgaben einen Abruf aus dem Gedächtnis erforderten oder in Form von Musterantworten lediglich zum Lesen aufforderten.

Aus der Kombination der beiden between-subjects Faktoren ergaben sich vier Typen von Aufgaben:

1) Abrufaufgaben, deren Beantwortung das Generieren von Zusammenhängen auf Basis multipler Informationen erforderte (Abrufübung mit Fokus auf Verständnis)

2) Musterantworten auf solche Abrufaufgaben lesen, die Zusammenhänge abdecken (Lesen mit Fokus auf Verständnis)

3) Abrufaufgaben, deren Beantwortung durch das Erinnern an isolierte Fakten möglich war (Abrufübung mit Fokus auf Wiedergabe)

4) Musterantworten auf solche Abrufaufgaben lesen, die isolierte Fakten abdecken (Lesen mit Fokus auf Wiedergabe)

Eine Woche später fand ein Leistungstest statt, der sowohl die Behaltensleistung für isolierte Fakten als auch das Verständnis von Zusammenhängen erfasste.

Vorläufige Ergebnisse

Die vorläufigen Ergebnisse der MANOVA mit Messwiederholung zeigten, dass die Wiedergabe im Leistungstest nach Abrufaufgaben besser gelingt (unabhängig davon, ob deren Fokus auf Fakten oder Zusammenhänge gerichtet war); für das Verständnis von Zusammenhängen im Leistungstest machte es dabei keinen Unterschied, ob man mittels Abruf- oder Leseaufgaben gelernt hat, $F(1, 186) = 6.85$, $p = .01$, $\eta p2 = .04$. Verständnis von Zusammenhängen im Leistungstest profitierte von Aufgaben, die den Fokus auf Verständnis legten (unabhängig davon, ob es Abruf- oder Leseaufgaben waren); wohingegen dies für die Wiedergabe keine Rolle spielte, $F(1, 186) = 7.49$, $p = .007$, $\eta p2 = .04$.

Diskussion

Um langfristig Verständnis zu fördern, ist es in der Lernphase empfehlenswert, den Fokus auf das Verständnis von Zusammenhängen zu legen. Dabei ist es irrelevant, ob dies über Abrufaufgaben erfolgt oder über Aufgaben, die keinen Abruf erfordern (wie Lesen von Musterantworten). Um langfristiges Behalten von Fakten zu unterstützen, sind Abrufaufgaben zu empfehlen. Dabei ist es irrelevant, ob die Abrufaufgaben isolierte Fakten betreffen (Fokus auf Wiedergabe) oder auf Zusammenhänge abzielen (Fokus auf Verständnis).

Unter dem Strich ist es empfehlenswert, eine Abrufübung mit Fokus auf Verständnis einzusetzen, weil damit sowohl das Verständnis als auch die Wiedergabe gefördert werden.

E-12: Von der Konzeptualisierung zur Messung: Instrumente zur Erfassung politischen Wissens und deren Anwendung in der empirischen Bildungsforschung

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* T03 R02 D26

Von der Konzeptualisierung zur Messung: Instrumente zur Erfassung politischen Wissens und deren Anwendung in der empirischen Bildungsforschung

Chair(s): **Pascal Alscher** (TU Dortmund), **Ulrich Ludewig** (TU Dortmund), **Nele McElvany** (TU Dortmund)
*Diskutant*in(nen):* **Jörn Sparfeldt** (Universität des Saarlandes)

Eine wichtige Voraussetzung für politische Beteiligung ist die Fähigkeit, politische Strukturen, Prozesse und Entscheidungen verstehen und bewerten zu können. Eine besondere Rolle spielt in diesem Zusammenhang das politische Wissen. Schon basale Partizipationsformen, wie das Wählengehen, setzen ein gewisses Maß an politischem Wissen voraus, wenn eigene Überzeugungen in politische Präferenzen und Wahlentscheidungen überführt werden sollen (Schübel, 2018). Verschiedene Studien haben gezeigt, dass Menschen, die viel über Politik wissen, vergleichsweise häufiger an Wahlen teilnehmen als Menschen, die wenig über Politik wissen (Delli & Carpini, 1996; Braun & Tausendpfund, 2019). Problematisch ist dies vor allem dann, wenn sich das Wissensniveau in unterschiedlichen Bevölkerungsgruppen systematisch unterscheidet. Erste interindividuelle Unterschiede entlang sozialer Differenzlinien werden auch in Deutschland bereits im Jugendalter festgestellt (z.B. Alscher et al., 2022; Hahn-Laudenberg & Abs, 2017). Bisher wenig diskutiert und dennoch von besonderer Bedeutung ist die Frage, wie politisches Wissen erfasst wird. Hierbei ist vor allem zwischen objektivem und subjektivem und zwischen theoretischem und handlungsbezogenem Wissen zu unterscheiden. Auch die Auswahl oder das Auslassen bestimmter Wissensbereiche kann problematisch sein (Oberle, 2012).

Im Rahmen des Symposiums sollen deshalb vor allem die verschiedenen Möglichkeiten der Erfassung politischen Wissens aufgezeigt und kritisch analysiert werden. Im Vordergrund steht die Auseinandersetzung mit bereits existierenden Messinstrumenten, deren Zusammenspiel mit weiteren Aspekten der politischen Sozialisation sowie ein Austausch zu etwaigen Weiterentwicklungspotentialen von Testinstrumenten für die zukünftige, systematische Erfassung politischen Wissens bei Schüler*innen.

Im ersten Vortrag wird ein Systematic Review zu politischen Wissenstests vorgestellt, in dem die Anforderungsstrukturen existierender Messinstrumente fokussiert werden. Hierfür wurden insgesamt 28 nationale und internationale Messinstrumente analysiert.

Im zweiten Vortrag wird das politische Wissen als Teil eines Civic Literacy Frameworks betrachtet und untersucht. Im Fokus des Beitrags liegt die Analyse der faktoriellen Struktur des theoretisch angenommenen Frameworks und die Betrachtung von Unterschieden zwischen Schüler*innen der siebten und zehnten Klassenstufe.

Der dritte Beitrag fokussiert auf etwaige Unterschiede im subjektiven und objektiven Wissen über die Europäische Union von Schüler*innen. Als Datengrundlage dient die wissenschaftliche Begleitung von Planspielen, im Rahmen derer das EU-Wissen von Schüler*innen in einem Prä-Post-Design erfasst wurde. Im Vortrag wird zudem der Frage nachgegangen, wie die Veränderungen im subjektiven und objektiven politischen Wissen mit Veränderungen im politischen Effektivitätsgefühl zusammenhängen.

Im vierten Beitrag werden Erwachsenendaten aus dem ALLBUS zum politischen Wissen betrachtet. Im Vordergrund stehen dabei die Fragen, was und wieviel Menschen über Politik wissen, wie sich das Wissensniveau in unterschiedlichen Bevölkerungsschichten unterscheidet und wie das Wissensniveau mit politischen Einstellungen und politischer Beteiligung zusammenhängt.

Beiträge des Symposiums

Wissensarten in Tests zum politischen Wissen: Ein Systematic Review

Rukiye Ates, Hermann J. Abs
Universität Duisburg-Essen

Theoretischer Hintergrund

Eine repräsentative Demokratie benötigt politisch kompetente Bürger*innen, damit sie funktionieren kann (Delli Carpini & Keeter, 1996, S.1). Politische Kompetenz basiert auf politischem Wissen, welches bei Bürger*innen vergleichsweise niedrig und ungleich verteilt ist (Converse, 2000, S.2). Allerdings sind die Definition und Operationalisierung des politischen Wissens strittig (Lawrence, 2003). Delli Carpini und Keeter (1996, S.10) definieren politisches Wissen als „the range of factual information about politics that is stored in long-term memory “. Neumann (1986) hinterfragt die Definition des politischen Faktenwissens und differenziert zwischen Faktenwissen und konzeptuellem Wissen. Converse (2000) unterstützt ebenfalls den Gedanken, dass Wissenstests sorgfältig konstruiert und aus mehreren Elementen bestehen sollten. In der Gegenwart warnen einige Forscher*innen davor, dass eine Faktenwissensabfrage, welche eher eine simple Wissensreproduktion darstellt, keine valide Messung des politischen Wissens ist (Stehr, 2015). Schübel (2017) vertritt hingegen die Position, dass durch Faktenabfrage ein allgemeiner Ist-Zustand zum politischen Wissen erfasst werden könne, da das erlernte Wissen im Gedächtnis dauernd in eine Art Netzwerkstruktur eingebettet existiere. Empirische Ergebnisse zeigen, dass eine politische Faktenwissensabfrage tatsächlich mehr repräsentieren kann als zuvor angenommen (Westle, 2005; Schübel, 2018).

Forschende sind zwangsläufig mit dieser Kontroverse konfrontiert, wenn sie bereits veröffentlichte Messinstrumente zum politischen Wissen für eigene Forschungsprojekte nutzen und adaptieren wollen. Auch bei der Entwicklung neuer Messinstrumente zur Erfassung des politischen Wissens stellen sich die Frage, welche Wissensarten berücksichtigt werden sollten. Dazu benötigt die Forschung einen Überblick über bereits veröffentlichte Messinstrumente mit Informationen zu deren Operationalisierung von politischem Wissen. Dieses kann anhand einer Itemkategorisierung innerhalb einer Wissenstaxonomie, wie von Marzano & Kandell (2007), welche die rein kognitive Lernzieltaxonomie von Bloom et al. (1956) ergänzt, durchgeführt werden. Zudem definieren sie Standards für Bildung neu und bieten eine Ressource für die Bildungsforschung. Marzanos & Kandells zweidimensionale Taxonomie mit sechs Leveln mentaler Prozesser innerhalb einer Dimension sowie drei Wissensbereichen in der anderen Dimension eignen sich optimal für die Kategorisierung von Wissensarten in den Messinstrumenten.

Fragestellung

Dieser Beitrag präsentiert die Ergebnisse eines Systematic Reviews internationaler und nationaler standardisierter Messinstrumente zur Erfassung politischen Wissens. Angesicht dessen soll der Frage nachgegangen werden: Welche Wissensarten des politischen Wissens werden in den Messinstrumenten erfasst?

Methode

Für die systematische Recherche wurden im Untersuchungszeitraum Mai 2021– März 2022, die Datenbanken Psyndex, ERIC, ERC, FIS und DAS überprüft. Insgesamt führte die Suche zu 800 Ergebnissen. Unter Berücksichtigung verschiedener Ausschlusskriterien (Suchbegriffe, Studiendesign, Erhebungsmethode, Stichprobegröße) wurden insgesamt N = 44 Messinstrumente gesichtet. Verfügbar waren letztlich N = 28 Messinstrumente (davon N = 9 deutsche; N = 14 internationale und N = 5 international vergleichende), die in die Analyse eingeschlossen wurden. Der Prozess wurde mit einer Literaturdatenbank in CITAVI (Version 6) und MS EXCEL (Version 2019) dokumentiert. Die ausgewählten Messinstrumente wurden angelehnt an die Wissenstaxonomie von Marzano und Kandell (2007) kategorisiert. Für die Auswertung der 28 Messinstrumente wurden die Level des kognitiven Systems 1 bis 4 verwendet: (1) Abrufen, (2) Verstehen, (3) Analyse, (4) Wissensnutzung. Insgesamt sind 1050 Items von zwei Rater*innen im Programm MaxQda (2022) kategorisiert worden.

Ergebnisse

Die Ergebnisse, welche mit einer Intercoder-Reliabilität von .80 nach Brennan und Prediger eine gute Zuverlässigkeit aufweisen, zeigen, dass 58,6 % der Items sich der Komponente Abrufen zuordnen lassen. Weitere 32,8 % der Items können der Kategorie Wissensnutzung zugeordnet werden. 7,4 % der Items erfordern eine Analyse und 1,3 % der Items beanspruchen das Verstehen. Insbesondere die deutschen Messinstrumente haben sich zu 94,5 % der Komponente Abrufen zuordnen lassen. Internationale Messinstrumente hingegen fordern neben 46% Abrufen, auch 37,4% Wissensnutzung und zu 16% das Analysieren. Die Ergebnisse zeigen insbesondere, dass die Testentwicklung im deutschsprachigen Raum von internationalen Entwicklungen profitieren kann.

Entwicklung und Erprobung eines Kompetenzmodels zur systematischen Analyse schulischer politscher Bildung

Pascal Alscher, Ulrich Ludewig, Nele McElvany
TU Dortmund

Theoretischer Hintergrund

Die Veröffentlichung der PISA 2000 Ergebnisse wurde in Deutschland von einer öffentlichen Empörung und Debatte über die vergleichsweise enttäuschenden Leistungen der deutschen Schüler*innen begleitet (Lange, 2002). Eine der Konsequenzen des ernüchternden Abschneidens im internationalen Vergleich war die Verabschiedung nationaler Bildungsstandards für eine Anzahl von Schulfächern. Dieser Schritt stellte einen Paradigmenwechsel weg von einer Input- hin zu einer Output-gesteuerten Bildungspraxis und –forschung dar (z. B. Abs & Klieme, 2005). Es wurden jedoch keine Standards für das Fach Politik verabschiedet, sodass es vor allem den Fachdidaktiker*innen sowie den Fachverbänden selbst oblag, geeignete Kompetenzmodelle zu entwickeln (Köller, 2016). Im deutschsprachigen Kontext ist vor allem das Modell der Politikkompetenz von Detjen et al. (2012)

bekannt. International sind zudem die internationalen Civic Education Studien sowie das in den USA durchgeführte Civics Assessment von Bedeutung.

Es fehlt jedoch ein Modell, welches die Eigenheiten des politischen Systems Deutschlands berücksichtigt, fachdidaktisch begründet ist und entsprechende Instrumente zur (längsschnittlichen) empirischen Erfassung enthält. Vor diesem Hintergrund wurde ein Civic Literacy Framework konzeptualisiert und operationalisiert, welcher einen politischen Wissenstest sowie weitere Dispositionen in den Bereichen Motivationen, Einstellungen und Volitionen enthält (Alscher et al., 2022). Zum Bereich der Motivationen zählen das politische Interesse und das politische Effektivitätsgefühl. Im Bereich der Einstellungen geht es um die Zustimmung zur Demokratie und die Zustimmung zur pluralen Gesellschaft. Für den Bereich der Volitionen wird zwischen einer zivilgesellschaftlichen und politischen Partizipationsbereitschaft unterschieden.

Fragestellung

Vor dem Hintergrund von Erhebungen in der siebten und zehnten Klassenstufe soll die Eignung des Frameworks für die Sekundarstufe I geprüft werden. Hieraus ergeben sich die folgenden beiden Forschungsfragen:

(1) Welche zentralen Strukturen einer Civic Literacy lassen sich auf Basis der theoretischen Vorannahmen empirisch stützen?

2) Welche Klassenstufenunterschiede lassen sich für die Civic Literacy und insbesondere das politische Wissen beobachten?

Methoden

Insgesamt haben N = 1.047 Schüler*innen an der Studie zur Entwicklung politischer und gesellschaftlicher Kompetenz im Jugendalter (EPKO) teilgenommen. Zum Zeitpunkt der Datenerhebung waren n = 613 Schüler*innen in der siebten (Malter = 12,47; 47,4% weiblich; 44,4% mit Migrationshintergrund) und n = 434 in der zehnten Klassenstufe (Malter = 15,49; 47,7% weiblich; 44,4% mit Migrationshintergrund). Zur Beantwortung der ersten Forschungsfrage wurden konfirmatorische Faktorenanalysen in RStudio mit dem „lavaan“ Paket gerechnet. Im Rahmen der zweiten Forschungsfrage wurde die Messinvarianz zwischen den beiden Klassenstufen getestet. Zudem wurden Mittelwertunterschiede zwischen den Klassenstufen sieben und zehn berechnet und auf Signifikanz überprüft.

Ergebnisse

Die Ergebnisse zeigen, dass die theoretisch angenommenen Strukturen des Frameworks weitgehend von den empirischen Daten gestützt werden. Eine Ausnahme ist das politische Effektivitätsgefühl, für das die Analysen eine Unterscheidung in ein internes und externes Effektivitätsgefühl nahelegen. Außerdem kann die theoretisch angenommene Unterscheidung in eine zivilgesellschaftliche und politische Partizipationsbereitschaft (Ekman & Amnå, 2012) nicht gestützt werden.

Eine Überprüfung der Verknüpfung zwischen Items und Konstrukten in der siebten und der zehnten Klassenstufe zeigt, dass Messinvarianz für die beiden Klassenstufen gegeben ist. Die Mittelwerte aller Konstrukte liegen in der zehnten Klassenstufe höher als in der siebten Klassenstufe. Statistisch

signifikante Mittelwertunterschiede wurden für das politische Wissen, das politische Interesse, die Einstellung zur Demokratie und die Einstellung zur pluralen Gesellschaft festgestellt. Dies deutet darauf hin, dass stärker partizipationsbezogene Aspekte einer Civic Literacy weniger entwicklungssensitiv sind. Ein Grund hierfür könnte sein, dass Partizipationsbereitschaft und Effektivitätsgefühl nicht nur von den eigenen Fähigkeiten und Interessen abhängen, sondern auch vom Vergleich mit den Peers bedingt werden.

Der Civic Literacy Framework soll künftig eine umfassende und systematische Erfassung von Kompetenz im Kontext der politischen Bildung ermöglichen und zielt darauf ab, Bildungsmonitoring und -evaluation voranzutreiben, um Forschung und Praxis der politischen Bildung zu fördern.

Erfassung von EU-bezogenen Wissenszuwächsen im Politikunterricht – eine Mixed-Methods-Untersuchung

Johanna Leunig, Monika Oberle
Georg-August-Universität Göttingen

Theoretischer Hintergrund

Politisches Wissen und politisches Effektivitätsgefühl sind zentrale Dimensionen der Politikkompetenz (Detjen et al., 2012) und damit relevante Ziele politischer Bildung. In den letzten Jahrzehnten haben Kompetenzen bezüglich der Europäischen Union (EU) als Ziel politischer Bildung in Deutschland und Europa an Bedeutung gewonnen (KMK, 2020). Um Ausgangslagen und Wirkungen politischer Bildung empirisch zu erfassen, bedarf es geeigneter Messinstrumente. Ein gängiger Weg der Operationalisierung politischen Wissens sind Batterien aus Single-Choice-Items zur Erfassung deklarativen Wissens, wobei erhobenes Faktenwissen als Proxy-Indikator für weitergehendes konzeptuelles Wissen angesehen wird (Delli Carpini & Keeter, 1996; Oberle, 2012). Auch Likert-skalierte Items und Item-Batterien zur Erhebung des subjektiven politischen Wissens werden häufig als Proxy-Indikator für objektives Wissen genutzt, was allerdings insbesondere für heterogene (z.B. schulformübergreifende) Untersuchungsgruppen problematisiert wurde (ebd.). Empirische Studien konnten positive Wirkungen von EU-bezogenen Bildungsinterventionen auf objektive und subjektive EU-Kenntnisse sowie auf das politische Effektivitätsgefühl von Schüler*innen nachweisen (Oberle & Leunig, 2018, Oberle et al., 2018).

Fragestellung

Der Beitrag geht der Frage nach, welche im Politikunterricht neu gewonnenen EU-Kenntnisse aus Sicht der Lernenden relevant erscheinen und mit einer Veränderung ihres politischen Effektivitätsgefühls einhergehen.

Methode

In einer Interventionsstudie wurden europapolitische Planspiele an 12 Schulen (Gymnasien, Gesamtschulen und Berufsschulen/Berufskollegs) begleitend beforscht und dabei ihre Wirkungen mit einem Mixed-Methods-Ansatz untersucht. Neben einer schriftlichen Erhebung mittels teilstandardisiertem Fragebogen im Prä-Post-Design (N=308, 51,2 % Mädchen; 57,8 % Gymnasiast*innen; MAlter=16,75 Jahre) wurden qualitative Interviews mit Teilnehmenden durchgeführt (N=12). Die quantitativen Analysen wurden überwiegend in SPSS 28 und latent in Mplus 8.8 (CFA, multiple Regressionen, SEM, LCM) vorgenommen. Das objektive EU-bezogene Wissen

wurde mittels 24 Single-Choice-Items (weiterentwickelt aus Oberle 2012, wo Expertenreviews die curriculare Validität bestätigt hatten) mit zumeist drei Distraktoren erhoben (Rasch-Skalierung in ConQuest). Im Mittelpunkt der leitfadengestützten Interviews stehen die von den Teilnehmenden wahrgenommenen Lerneffekte der EU-Simulation. Die Antworten der Schüler*innen wurden mittels einer computergestützten kategorialen Inhaltsanalyse (Software Maxqda) ausgewertet und können mit ihren Fragebogenantworten verknüpft werden.

Ergebnisse

Die Prä-Post-Mittelwertvergleiche bestätigen bedeutsame Zuwächse des objektiven und subjektiven EU-Wissens sowie des politischen Effektivitätsgefühls der Teilnehmenden (Cohens d=.42 bis .58). Latent-Change-Modellierungen zeigen, dass objektive und subjektive Wissenszuwächse nur moderat korreliert sind und dass die Veränderung im externen politischen Effektivitätsgefühl (Responsivitätsgefühl) maßgeblich durch die subjektiven, jedoch nicht durch die objektiven Wissenszuwächse erklärt wird. Die Auswertung der qualitativen Interviews ergibt, dass die Schüler*innen Lerneffekte in den Dimensionen polity (Form) und policy (Inhalt), insbesondere jedoch politics (Prozess) berichten und dabei besonders gewonnene „Einsichten“ (Fischer, 1993) hinsichtlich der Komplexität von Entscheidungsprozessen sowie der Relevanz und Schwierigkeit von Kompromissfindungen in der Politik betonen. Abschließend soll diskutiert werden, inwiefern diese subjektiv wahrgenommenen Lerneffekte bei der Operationalisierung von objektivem politischen (EU-)Wissen besser berücksichtigt werden können.

Politisches Wissen in Deutschland: Niveau, Determinanten und Konsequenzen

Markus Tausendpfund
FernUniversität in Hagen

Theoretischer Hintergrund

Demokratie ist eine anspruchsvolle Regierungsform. Entscheidungsprozesse sind häufig komplex und Verantwortlichkeiten nicht immer klar erkennbar. Ein aufgeklärtes Politikverständnis („enlightened understanding“) der Bürgerschaft gilt deshalb als eine zentrale Voraussetzung für das Funktionieren einer Demokratie (Dahl, 2000, S. 37-39). Die Menschen sollten wissen, wie das politische System funktioniert und nach welchen Prinzipien Entscheidungen getroffen werden. Um eine begründete Entscheidung für oder gegen eine Partei, einen Kandidaten oder eine Sachfrage treffen zu können, ist die Kenntnis der zentralen Akteure sowie ihrer Vorschläge und Standpunkte von zentraler Bedeutung (Delli Carpini & Keeter, 1996).

Während insbesondere für die USA umfangreiche Forschungsbefunde zu Niveau, Determinanten und Konsequenzen politischer Kenntnisse vorliegen, sind die empirischen Befunde für die erwachsene Bevölkerung in Deutschland überschaubarer (Tausendpfund & Westle, 2020). In der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) wurde 2018 (erneut nach 2008) ein größeres Politik-Quiz zur Erfassung politischen Wissens realisiert, sodass eine aktuelle Datengrundlage existiert, um das Wissensniveau der Bevölkerung zu beschreiben, zentrale Bestimmungsfaktoren zu identifizieren und Konsequenzen des politischen Wissens zu untersuchen. Insbesondere bieten die Daten die Möglichkeit, Stabilität und Veränderung der Wissensstände zu zwei Zeitpunkten analysieren zu können.

Fragestellung

Auf Basis der ALLBUS-Daten wird zunächst das politische Wissensniveau der Bevölkerung zu zwei Zeitpunkten (2008 und 2018) beschrieben und zentrale Determinanten des politischen Wissens identifiziert. Dabei werden drei Fragen beantwortet: Was und wie viel wissen die Menschen in Deutschland über Politik? Wie unterscheidet sich das Wissensniveau in verschiedenen Bevölkerungsgruppen? Und: Was sind Folgen des politischen Wissensniveaus für Einstellungen gegenüber Politik und politischer Beteiligung?

Methoden

Mittels einer Mokkenanalyse wird zunächst geprüft, ob den Wissensitems ein eindimensionales latentes Konstrukt – hier: politisches Wissen – zu Grunde liegt (Moosdorf et al., 2020). Daran anknüpfend wird auf der Basis von linearen und logistischen Regressionsmodellen zunächst der Einfluss individueller Merkmale auf das politische Wissensniveau untersucht, ehe Zusammenhänge zwischen dem politischen Wissen und politischen Einstellungen bzw. politischem Verhalten untersucht wird.

Ergebnisse

Die empirischen Befunde zeigen, dass das durchschnittliche Wissensniveau der Bürger*innen zwar nicht dem Idealbild des „homo politicus“ entspricht, aber die meisten Menschen über solide Kenntnisse des politischen Systems verfügen. Allerdings ist das politische Wissensniveau in der Bevölkerung höchst ungleich verteilt. Eine höhere Bildung, eine höhere Schichtzugehörigkeit, ein höheres Alter sowie ein stärkerer Nachrichtenkonsum der öffentlich-rechtlichen Fernsehsender korrelieren mit einem höheren Wissensniveau. Dagegen hängen Migrationshintergrund und das weibliche Geschlecht mit geringerem politischem Wissen zusammen.

Die soziale Ungleichheit beim politischen Wissen wäre allerdings nicht problematisch, wenn das politische Wissen keine Folgen für politische Einstellungen und Verhalten hätte. Die empirischen Befunde widersprechen allerdings dieser Annahme. Mit einem höheren Wissensniveau zeigt sich nicht nur eine stärkere Unterstützung der Demokratie als Regierungsform, sondern auch eine stärkere Bereitschaft sich politisch zu beteiligen. Deshalb kann politisches Wissen auch in Deutschland als „cornerstone construct in research on political behavior“ (Mondak, 2001, S. 238) bezeichnet werden.

E-13: A Multi-Faceted Perspective on Current Approaches to Reflect on Teacher Reflections

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* T03 R02 D39

A Multi-Faceted Perspective on Current Approaches to Reflect on Teacher Reflections

Chair(s): **Tim Fütterer** (University of Tübingen, Germany)
*Diskutant*in(nen):* **Margarete Imhof** (Johannes Gutenberg University Mainz, Germany)

Worldwide, the ability to reflect on teaching practices is a key characteristic of professional teachers and thus an aim of teacher training programs (Darling-Hammond, 2021; Korthagen, 2014). However, it is still discussed, how to conceptualize teachers' reflective competence and how to assess as well as to promote teachers' reflections. Therefore, in this symposium we bring together current approaches from different labs that aim to conceptualize reflective competencies of teachers, to assess teachers' reflections reliable, valid, and efficiently (e.g., AI-based), and to promote reflections systematically.

The first paper focuses on the conceptualization of teacher reflective competence by addressing its relationship to other teacher competencies. Results of a conceptual literature analysis concerning the depth and content of reflective processes lead to the conclusion that reflection can be understood both as a skill that is needed for the development of other competences as well as an independent competence whose development should be promoted.

To assess and promote reflective processes of teachers written reflections (e.g., portfolios: Fütterer, 2019) are often used. However, validated instruments to assess the quality of written reflections are scarce. Therefore, in the second paper an instrument to assess the quality of written reflections of (pre-service) teachers is introduced. The quality of reflections is operationalized using theoretical approaches regarding the concept of professional vision. First results from a comprehensive experimental study indicate that the instrument is reliable and valid.

In the third paper the authors refer to a model that describes reflections on different quality levels to analyze written reflections of pre-service teachers. The results of structured qualitative content analyses indicate that pre-service teachers' written reflections are rather at low and middle but not at high levels. Linguistic features (e.g., semantics) were not statistically different between levels of reflection. Both Paper 3 and 4 provide an outlook on how the current work will be used for AI-based assessments of written reflections.

In the fourth paper the authors focus on the question of how written reflections of pre-service teachers can be systematically stimulated. Based on theoretical assumptions about social agency an experimental approach is used by varying the addressees of reflections. The results illustrate that the addressee (i.e., peer vs. teacher) makes a difference in how pre-service teachers frame their written reflections; visible in text characteristics (e.g., number of words) and linguistic features (e.g., coherence).

A distinguished discussant with substantive expertise in teacher reflective processes will discuss all contributions.

Beiträge des Symposiums

An Independent Competence Area or Mother of All Competences: The Role of Reflectivity in a Teacher's Competence

Minna Körkkö[1], Laura Ketonen[2]
[1]University of Oulu, Finland, [2]University of Jyväskylä, Finland

Theoretical Background. Reflection is a key facilitator in teacher learning. The starting point of this study is our notion that in the context of teacher education reflection is often viewed as a tool for professional development, with which the teachers develop their own work and themself as teachers (Körkkö, 2020; Toom et al., 2015). However, this kind of definition does not describe how reflection actually relates to the teacher's transversal competence and how reflection can be perceived as part of the development of teaching. In addition, reflection is also known to be related to other educational practices, not just to the development of teaching, which also advocates a more versatile examination of the concept (Beauchamp, 2014). In this study, the authors discuss reflectivity as a part of a teacher's competence by utilizing a multidimensional adapted process model of teaching (MAP model; Blömeke et al., 2015; Metsäpelto et al., 2021). The starting point of the study is our notion that teachers' reflection skills are not explicitly present in the model but are described as part of other skills. In the model, reflection is understood quite narrowly as a cognitive skill related to critical thinking, even though research has shown the many meanings and usefulness of the concept of reflection in several different contexts, for example in assessing students' skills (Kember et al., 2008). Considering how important reflection is for teacher work and development, and how important the MAP model has become in teacher education, we see it important to investigate the relationship of the concept of reflection to other areas of teacher competence presented in the MAP model.

Research Questions. The research was guided by the following research questions:

(RQ1) How is reflection described in the literature of teacher competence?

(RQ2) How does the reflection described in the literature relate to the teacher's competence areas?

As the result of the conceptual analysis, we present how reflection could be related to the teacher's competence, possibly as an independent area of competence or as a broader general framework of competence.

Method. The analysis includes two levels, in the first of which we analyzed the literature about areas of teacher competence, especially the parts dealing with reflection. We analyzed the depth of reflection using Mezirow's (1981, 1991) transformative learning theory and the contents of reflection using Korthagen's (2017) holistic perspective. We used the four-step reflection classification developed and tested by Kember and colleagues (2000, 2008; see also Peltier et al. 2005), especially based on Mezirow's thinking, where reflection is viewed as a continuum from non-reflective to reflective activity. Classification is used in this study as an analysis tool of Mezirow's concept of reflection. At the second level of analysis, we examined the relationship of reflection described in the literature and the teacher's areas of competence.

Results. Reflection is connected to all areas of a teacher's competence through situation-specific skills (perception, interpretation, decision-making) and professional practices. The descriptions of reflection in the articles included in the review almost completely lacked the level of intensive, critical reflection included in Mezirow's model. The various learning areas of Korthagen's model were widely

represented in the articles. The results indicate that reflection can be understood both as the "mother of all competences", i.e., as a skill that is needed for the development of other competence areas, and as an independent competence area whose development should be supported. The study discusses what this could mean in terms of the MAP model and, more broadly, the development of teacher education.

Hit the Road Towards an Automated Assessment of Written Reflections of Pre-Service Teachers

Tim Fütterer, Kathleen Stürmer, Richard Göllner
University of Tübingen, Germany

Theoretical Background. Videotaped teaching situations are a promising approach to promote reflective processes of student teachers (Gaudin & Chaliès, 2015). Videos are suitable because they represent deliberate and targeted reductions of the complexity of teaching situations, especially helpful for learning processes of novices (see Grossman et al., 2009). A common method for obtaining data and thus gaining insight into teachers' reflective processes on instructional videos is through written reports (Santagata et al., 2021). However, instruments to assess teacher reflections about teaching captured in written reports (“written reflections”) that have been systematically tested for their quality (i.e., reliability and validity) are scarce. Therefore, less is known about the extent to which instruments allow for the assessment of different reflection aspects. Moreover, it is unclear, whether the quality of instruments depends, for instance, on the type of segments (e.g., sentences or sense units) or on the specificity of the prompts used to stimulate reflective processes.

Research Questions. We will present an instrument to evaluate pre-service written reflections which is a comprehensive coding scheme we developed to assess the quality of written reflections based on the concept of reasoning (see theory of professional vision; e.g., Stürmer & Seidel, 2015). Reasoning encompasses the processes describing and explaining teaching situations as well as predicting possible consequences of different instructional events for students. We address the following research questions (RQ):

(RQ1) How reliably can segments and aspects of reasoning be assessed in written reflections with the instrument?

(RQ2) How is the reliability associated with context factors?

(RQ3) How valid is the instrument?

Method. N = 200 pre-service teachers were systematically prompted to reflect on one to four staged videos (duration of 2 to 4 minutes) of mathematics and language lessons. The videos are representative of two teaching quality indicators: goal clarification and feedback (Seidel & Stürmer, 2014). 624 written reflections available; each at least 400 words long. Four intensively trained assistants segmented (sentence, sense units) the written reflections and coded the segments regarding the construct reasoning. Written reflections are segmented twice, 25% of the segments are coded twice. So far, we analyzed n ≈ 30.000 segments (sentence). To test the reliability (RQ1), we calculated inter-rater agreements. We conducted t-tests and ANOVAS to assess differences of inter-rater agreements in context factors (raters, teaching quality indicators, prompts, and video stimuli

[RQ2]). To test the validity (RQ3), we will analyze correlations of the quality of written reflections and test data collected with the Observer Research Tool (Stürmer & Seidel, 2017).

Results. RQ1: The results indicate a high inter-rater agreement (M = 0.91, SD = 0.07) for sentence segmentation. These preliminary findings indicate that our instrument proofed to be reliable to assess sentence segments in written reflections. RQ2: The agreement was not different between rater pairs ($t(622) = 0.024$, $p = .981$, $d = 0.00$), nor teaching quality indicators ($t(622) = -0.846$, $p = .398$, $d = 0.00$), nor prompts of different specificity ($F(3, 620) = 2.07$, $p = .102$, $\eta = .010$) but between video stimuli ($F(4, 619) = 2.84$, $p = .024$, $\eta = .018$). By the end of 2022, all data will be segmented and coded. We will be able to present comprehensive results on the reliability (i.e., regarding coding of sense units and the quality of reflections) as well as on the validity of the instrument (RQ3). Furthermore, as artificial intelligence (AI) algorithms enable automated coding (Wang et al., 2022), we will outline how the inclusion of extensive information (i.e., rating information derived from variations in terms of segmentation and prompting) will be used to train an AI algorithm to automatically analyze written reflections and to understand which information are needed to train an AI algorithm to be reliable and valid.

Quality of Pre-Service Teachers' Reflections—A Mixed-Method Study

Chengming Zhang, Jessica Schießl, Lea Plößl, Florian Hofmann, Michaela Gläser-Zikuda
University of Erlangen-Nürnberg, Germany

Theoretical Background. In teacher education, reflection is a crucial aspect of professionalization (Hatton & Smith, 1995; Korthagen, 2001) and may be supported by a portfolio (Gläser-Zikuda et al., 2020). To analyze reflective writing and the quality of reflection, in recent years, researchers, on the one hand, used qualitative analyses (Fütterer, 2019; Poldner et al., 2014); on the other hand, automated and multi-dimensional approaches became relevant (Kovanović et al., 2018; Ullmann, 2019). However, both qualitative and quantitative studies have limitations. Qualitative research is time-consuming and labor-intensive (Jung & Wise, 2020), and quantitative research is difficult to generalize due to the lack of a theoretical basis (Cui et al., 2019). Therefore, applying theory-based data science methods in analyzing the quality of reflection should be valued. In this context, reference should also be made to the ever-increasing use of AI in education (Pedro & Subosa, 2019), which gives hope for analysis methods that significantly outperform previous automatic methods.

Research Questions. Therefore, the primary goal of the BMBF-funded project was to conduct interdisciplinary research on the utilization of machine learning methods in teacher education, particularly in natural language processing (NLP), and its implementation in artificial intelligence (AI) based portfolios. The study initially formulated the following three research questions:

(RQ1) On which level can the reflective writings of pre-service teachers on profession-related topics be classified?

(RQ2) How do the psycholinguistics of reflective writings differ among pre-service teachers at different levels of reflection?

(RQ3) How do pre-service teachers at different levels of reflection differ in the topics of their reflective writings?

Method. N = 709 pre-service students had written reflections on specific tasks using the E-Portfolio. During the winter semester of 2021, 1012 reflective writings were collected. Two hundred reflective writings from pedagogical diagnostic and classroom management were randomly chosen and anonymously. After removing the outliers, the subsequent analysis was performed utilizing 198 reflective writings. The average number of words in these reflective writings was 229.85, with a standard deviation of 148.99, ranging from a high of 806 to a low of 30. Furthermore, 71.90% samples were female students, and 87.83% of students were from the first semester to the third semester. For analysis, we applied a structured qualitative content analysis (Gläser-Zikuda et al., 2020) based on the theory of Hatton and Smith (1995), dictionary-based linguistic analysis software Linguistic Inquiry and Word Count（LIWC; Pennebaker et al., 2015) and the unsupervised learning algorithm BERTopic (Grootendorst, 2022).

Results. The results of the content analysis indicated that the hierarchy of pre-service teachers' reflections was distributed at the low and middle levels (Description = 37, Reflective Description = 129, and Dialogic Reflection = 31), while reflections at the high level (Critical Reflection = 1) were not too many. Second, statistical differences between reflection levels were demonstrated for 13 psycholinguistic attributes, particularly in the categories of cognition and affect. Finally, based on the BERT topic modeling, four topics were identified in the reflective writings: Learning materials (n = 94), motivation and emotions (n = 31), learning content (focus on pedagogical diagnostics) (n = 20), and learning content (focus on classroom management) (n = 15). The higher the level of reflection, the more the topics of reflective writings were focused on motivation and emotion (Description 2.70%, Reflective Description 17.83%, and Dialogic Reflection 22.58%). Central results and the selected analysis method are presented and discussed with a view to methodological challenges and implications for teacher education. In this context, the role AI methods can play in evaluating reflections and which feedback options can be generated for the students is also discussed.

How Does the Role of the Addressees Determine how one Reflects?

Felina Leng, Fütterer Tim, Kathleen Stürmer, Richard Göllner
University of Tübingen, Germany

Theoretical Background. The ability to reflect teaching is a key competence of teachers (Seidel et al, 2013). For stimulating the acquisition of reflections competencies, student teachers oftentimes reflect real classroom situations in which they have to notice (i.e., identify) and reason (e.g., describe, explain) critical situations (e.g., Gröschner et al., 2018; Santagata et al., 2021). Instructional science showed that prompting is an efficient way to support generative learning activities, for instance, when they ask learners to learn by self-explaining or explaining to someone else (Fiorella & Mayer, 2016). Generally, prompts are described as activators to support the elaboration of the learning content (i.e., connect and compare relevant concepts). In addition to these cognitive principles, prompts may foster an in-depth processing of the learning content because learners must explain the content to someone else (Lachner et al., 2021) and especially to someone who is expected to be more knowledgeable. In line with the social agency theory (Atkinson et al., 2005), providing social cues may prime social responses in learners that lead to deeper cognitive processing during learning and thus better

performance. Thereby, explanations to agents assuming to reveal a higher competency might be particularly effective.

Research Questions. In the present study, we investigated the potentially beneficial effect of social prompting on the quality of teacher reflections on classroom situations. By using an experimental within- and between-person variation of the addressee of reflections, we examined whether social prompting condition outperform the quality of reflections in contrast to a non-social condition. Specifically, our study was guided by the following research questions (RQ):

(RQ1) How does the quality of pre-service teachers' written reflections differ between written reflections with and without an addressee?

(RQ2) How does the quality of pre-service teachers' written reflections differ between written reflections with different competency levels of the addressees?

Method. In an experimental study, N = 115 pre-service teachers had to write written reflections on classroom situations represented in one of four real but staged videos. Experts rated the videos to be representative for two quality aspects of teaching: goal clarification and feedback (Seidel & Stürmer, 2014). Pre-service teachers were prompted randomly to either imagine they explain their reflection to a peer (N = 60) or to a teacher (N = 55). Thus, the addressee of the explanation varies systematically between the conditions (social prompting). N = 85 written reflections with non-social prompting functioned as the control condition. To operationalize the quality of written reflections, we used linguistic features (e.g., average length of words and sentences, coherence). To answer the RQs, we conducted t-tests.

Results. Regarding RQ1 findings were mixed. Whereas mean word count ($p = .010$, $d = 0.37$) and readability ($p < .001$, $d = 0.61$) were higher in the condition with addressees than in the control condition, mean sentence length ($p < .001$, $d = 0.80$) was smaller and coherence ($p = .021$, $d = 0.34$) was lower. Regarding RQ2, the condition with a teacher as the addressee of their reflection wrote significantly more words per text ($d = 0.38$, $p = .040$). If writing about goal clarification, reflections addressed to a teacher were descriptively more coherent than those addressed to a peer ($d = 0.55$, $p = .054$). An additional coding of the frequency of words related to the quality aspects of teaching is in progress. We will present results regarding these frequencies (indicating the correctness of the contents) and findings based on the instrument introduced in Paper 2. Furthermore, as mixed findings may result from using a control condition participants were not randomly assigned to, we will give an outlook on an improved follow-up experiment.

E-14: Potentiale von Log-Daten in authentischen computerbasierten Assessments

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* T03 R02 D81

Potentiale von Log-Daten in authentischen computerbasierten Assessments

Chair(s): **Ulf Kroehne** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany)
*Diskutant*in(nen):* **Stadler Matthias** (Ludwig-Maximilians-Universität München)

Computerbasiertes Assessment (CBA), wie es in der Bildungsforschung zunehmend nicht nur in groß angelegten Bildungsvergleichsstudien, sondern auch in einzelnen Forschungs-projekten Anwendung findet, verspricht eine Vielzahl von Potentialen. Während bei CBA zunächst auch papierbasiert administrierbare Aufgaben dominierten (bspw. zur Steigerung der Messeffizienz und bei der Erforschung von Modus-Effekten), treten nun zunehmend Forschungsfragen und Anwendungen in den Fokus, welche speziell für CBA entwickelte Instrumente und Aufgabenformate nutzen. Die in diesem Symposium zusammengestellten Beiträge zeichnen sich dadurch aus, dass sie in unterschiedlichen Phasen des Forschungsprozesses und aus unterschiedlicher disziplinärer Perspektive die diagnostische Nutzung von Potentialen digitaler Erhebungsformen adressieren. Eine spezifische Verbindung zwischen den Beiträgen wird dabei durch die theoriegeleitete Nutzung von Log-Daten hergestellt, welche in authentischen Instrumenten mit reichhaltigem und interaktivem Aufgabenmaterial gesammelt werden können.

Beitrag 1 berichtet die Entwicklung und psychometrische Modellierung eines simulations-basierten Testinstruments zur Messung der Fähigkeit zum Lösen technischer Probleme im Alltag. Die entwickelten Items erlauben auf Grundlage der Log-Daten unter anderem eine Trennung des Wissenserwerbs über simulierte technische Geräte durch Explorationsverhalten versus Instruktionslesen. Darüber hinaus wird der Zusammenhang der Testwerte mit metakognitiver Genauigkeit berichtet.

Beitrag 2 behandelt spezifisch die Erfassung von Metakognition ebenfalls bei einer Problem-löse¬aufgabe und nutzt dabei konkrete Verhaltensweisen als Indikatoren für metakognitive(s) Monitoring und Kontrolle, die aus den Log-Daten einer simulierten Train Track Task gewonnen und kodiert werden können. Die verhaltensbasierten Indikatoren werden dabei mit einer Metakognitionsmessung im Fremdbericht validiert.

Beitrag 3 fokussiert am Beispiel eines Instruments zur Ökonomischen Bildung authentische Assessments, wie sie sich durch die Integration von Audio-, Video- und Bildelementen in die Aufgaben- und Instruktionsgestaltung erstellen lassen. Dabei werden unter Berücksichtigung von Expertenratings in einer Validierungsstudie Zusammenhänge zwischen der Authentizität der Items und schwierigkeitserzeugenden Merkmalen untersucht.

Beitrag 4 analysiert die mit Hilfe einer computerbasierten Supermarktsimulation erfassten Indikatoren für Aufmerksamkeitsverhalten im Hinblick auf das Kaufverhalten von Kindern. Die Zusammenhänge verschiedener aus den Log-Daten extrahierten Indikatoren werden dabei mit Hilfe konfirmatorischer Faktorenanalysen modelliert und es werden Zusammenhänge zwischen Kontrollverhalten und Aufgabenerfolg berichtet.

In einer abschließenden Diskussion werden die Beiträge im Kontext der empirischen Bildungsforschung, pädagogischen Psychologie und empirischen Pädagogik kritisch gewürdigt und kommentiert.

Beiträge des Symposiums

Assessment der Fähigkeit zum Lösen technischer Probleme im Alltag

Marcus Schrickel[1], Jennifer Stemmann[2], Carolin Hahnel[3]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany, [2]Pädagogische Hochschule Freiburg, Institut für Chemie, Physik, Technik und ihre Didaktiken, [3]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany; Zentrum für internationale Bildungsvergleichsstudien (ZIB)

Theoretischer Hintergrund

Technische Geräte, wie Heizungsthermostate, Waschmaschinen und Telefone, begleiten uns im Alltag und bieten im Zuge ihrer Weiterentwicklung neue oder zusätzliche Handlungsmöglichkeiten. Allerdings sind erforderliche Handlungsprozeduren oft nicht routiniert, wodurch die Nutzung technischer Alltagsgeräte zum Problem werden kann (vgl. Dörner, 1976). Technische Probleme sind Situationen, in denen Personen mit einem unvertrauten technischen System interagieren, um es von einem Ausgangszustand unter Überwindung von Barrieren in einen gewünschten Zielzustand zu überführen. Sie erfordern Fähigkeiten von Nutzer:innen, sich aktiv und effizient Bedienwissen über sich ändernde Technik anzueignen und es zielführend einzusetzen. Wie Personen jedoch Bedienwissen im Rahmen des technischen Problemlösens (TPL; Stemmann, 2016) erwerben, ist bisher wenig systematisch erforscht.

In der Problemlöse-Literatur werden die Prozesse des Wissenserwerbs (Erkunden des Problemraums) und der Wissensanwendung (Bearbeiten einer konkreten Aufgabe) unterschieden (Funke, 2001), die unter anderem durch Merkmale eines Problems (z. B. Komplexität des Funktionsumfangs, Intransparenz der Benutzeroberfläche) beeinflusst werden. Beim TPL kann relevantes Bedienwissen durch Systemexploration (z. B. Van Der Linden et al., 2001) oder das Erlesen von Informationen aus Bedienungsmanualen (vgl. information problem solving; Brand-Gruwel et al., 2009) erworben werden. Dabei kann das Verhalten beim Wissenserwerb durch gerätespezifisches Vorwissen beeinflusst sein. Ein erfolgreicher Wissenserwerb sollte wiederum zu einer erfolgreichen Anwendung von Wissen führen (z. B. Wüstenberg et al., 2012). Ein weiterer Einfluss auf das TPL könnte auch eine akkurate Einschätzung der eigenen Problemlösefähigkeiten (metakognitive Genauigkeit; Mihalca & Mengelkamp, 2020) sein, die im Zusammenhang mit der Zuweisung kognitiver Ressourcen beim Problemlösen steht (Dentakos et al., 2019).

Fragestellung

Um das technische Problemlösen von Personen untersuchen zu können, haben wir ein simulationsbasiertes Testinstrument zur Abbildung von TPL-Fähigkeiten entwickelt. Die resultierenden Testwerte unterziehen wir einer empirischen Überprüfung und ersten Validierung. Dazu untersuchen wir die psychometrischen Eigenschaften der entwickelten Items (Fragestellung 1). Wir prüfen hierbei, ob die entwickelten Items dieselbe Fähigkeit erfordern und inwiefern die Merkmale eines technischen Geräts die Lösungswahrscheinlichkeit der Items beeinflussen. Des Weiteren untersuchen wir, wie das Generieren von Handlungswissen über technische Geräte mit dem TPL-

Testwert zusammenhängt (Fragestellung 2). Dazu betrachten wir, in welchem Ausmaß Personen Wissen über das Explorieren eines technischen Geräts oder das Lesen von Instruktionen erwerben und welche Rolle gerätebezogenes Vorwissen hierbei einnimmt. Nicht zuletzt untersuchen wir den Zusammenhang vom TPL-Testwert und der metakognitiven Genauigkeit (Fragestellung 3).

Methode

Zur Erfassung des TPL wurden neun Simulationsumgebungen von Stemmann (2016) zu einem Test mit 36 Items weiterentwickelt. Eine TPL-Simulation präsentiert zunächst ein technisches Gerät (z. B. ein digitales Heizungsthermostat) und erfragt das gerätespezifische Vorwissen von Probanden, die das Gerät anschließend selbstständig erkunden und Funktionen in einem Bedienungsmanual nachschlagen können. Nach Abschluss dieser Wissenserwerbsphase werden den Probanden vier Items vorgelegt, in denen sie eine bestimmte Handlung am Gerät vornehmen müssen (z. B. „Stellen Sie das Heizungsthermostat so ein, dass die Tagtemperatur 21°C und die Nachttemperatur 17°C beträgt.“). Bevor sie die Items bearbeiten, werden sie gebeten einzuschätzen, wie viele Aufgaben sie korrekt lösen werden. Anhand von Logdaten wird anschließend bestimmt, ob, wie lange und wie umfassend Personen ein technisches Gerät in der Wissenserwerbsphase erkundet und/oder das Bedienungsmanual genutzt haben.

Zur Bearbeitung der Forschungsfragen wird die Erhebung einer Stichprobe von n = 280 Erwachsenen aus der breiten Bevölkerung angestrebt, wobei alle Items bearbeitet werden (minimally balanced repeated treatment design; Frey et al., 2009). Neben den Leistungs- und Prozessdaten aus dem TPL-Test werden auch demografische Angaben und das Immersionserleben (angelehnt an Iachini et al., 2019) erfasst.

Ergebnisse und deren Bedeutung

Die Datenerhebung findet gegenwärtig statt. Die Ergebnisse aus Analysen mit IRT- (bzgl. Dimensionalität) und GLMM-Modellen (bzgl. schwierigkeitsbestimmender Merkmale, Wissenserwerb, metakognitiver Genauigkeit) werden zum Zeitpunkt der GEB 2023 vorliegen und präsentiert. Sie dienen unter anderem der Diskussion der Gemeinsamkeiten und Unterschiede von TPL und anderen Formen des Problemlösens.

Technologiebasierte Erfassung von Metakognition im Zuge einer Online-Problemlöseaufgabe bei jungen Schulkindern

Bernadette van Berk[1], Donna Bryce[2], Ulf Kroehne[3], Charlotte Dignath[1]
[1]Institut für Schulentwicklungsforschung, TU Dortmund, [2]Universität Tübingen, [3]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany

Theorie

Metakognition – das Ausüben von metakognitivem Monitoring und metakognitiver Kontrolle – ist zentraler Bestandteil für selbstreguliertes Lernen und damit eine wichtige Voraussetzung für Lernerfolg. Eine große Herausforderung für die Forschung zu Metakognition stellt ihre Erfassung dar. Bisherige Studien zeigen, dass Personen nur unzureichend Auskunft geben können über ihre metakognitiven Aktivitäten, und dass sie ihren Einsatz von metakognitiven Strategien retrospektiv überschätzen. Zudem weisen Selbstberichtsmaße nur geringe bis keine Korrelationen mit Leistungsmaßen auf (Artelt, 2000; Veenman et al., 2006; Veeman & Van Cleef, 2019), weshalb

Selbstberichtsmaße keine validen Messinstrumente darstellen. Üblicherweise dienen Selbstberichtsmaße wie Fragebögen dazu, eine allgemeine Disposition für Metakognition zu messen. Neuere Befunde zeigen jedoch, dass Metakognition in Abhängigkeit von der Lernsituation intraindividuell stark variieren kann und darum im Prozess erfasst werden sollte (Winne & Perry, 2000).

Dennoch wird ein Großteil der Metakognitionsforschung nach wie vor mithilfe meist retrospektiver Fragebögen erfasst. Neben dem allgemeinen Problem der Überschätzung, möglichen Verzerrungen der Erinnerung, und eventuell sozialer Erwünschtheit spielt beim Selbstbericht zudem die Fähigkeit zur Introspektion, des sprachlichen Verständnisses sowie der sprachlichen Ausdrucksfähigkeit eine wesentliche Rolle für die Validität der Messung (Lockl et al., 2018). Dementsprechend stellt die Messung metakognitiver Aktivitäten bei jungen Lernenden eine zusätzliche Herausforderung dar.

Ein innovativer Weg, um Metakognition zu erfassen, stellt das Auswerten von Logfile-Daten dar. Dabei werden konkrete Verhaltensweisen als Indikatoren für Metakognition identifiziert und operationalisiert, die anhand der Logfiles prozessual erfasst werden können, um Darstellungen des Lernens im Verlauf zu erhalten. Lernen wird dabei als sukzessive Abfolge von Zuständen interpretiert, mit einem klar definierten Start- und Endpunkt jedes Zustands (Winne & Baker, 2013). Zur Validierung der Logfile-Daten können diese Indikatoren mit anderen Metakognitionsmaßen zusammengebracht werden.

Bislang liegt nur wenig Forschung zur Messung metakognitiver Aktivität von jungen Lernenden mithilfe von Logfile-Daten vor. In der vorliegenden Studie validieren wir die aus der Theorie hergeleiteten Indikatoren für metakognitive(s) Monitoring und Kontrolle gemessen mit Logfile-Daten beim Bearbeiten einer Problemlöseaufgabe und vergleichen diese mit einer Metakognitionsmessung im Fremdbericht (Elternfragebogen)an einer Stichprobe von Schulkindern der 2. bis 5. Jahrgangsstufe.

Methode

An der Online-Studie nahmen 80 Kinder teil. Zur Messung der Metakognition bearbeiteten die Teilnehmenden die Train Track Task (TTT; Brzce & Whitebread, 2012), eine Problemlöseaufgabe, bei der Schienenteile zu einem vorgegebenen Schienenkreis zusammengesetzt werden müssen. Dabei wird die Metakognition während des Bearbeitens der Aufgabe mithilfe eines systematischen Beobachtungssystems kodiert. In unserer Studie wurde die TTT im CBA ItemBuilder (Rölke, 2012) erstmals als online zu bearbeitende Aufgabe umgesetzt, und es wurden Logfile-Daten erhoben, die als Indikatoren für Metakognition während der Aufgabenbearbeitung genutzt wurden.

Auswertung

Im ersten Auswertungsschritt wurde ein Kodierschema für Indikatoren von metakognitivem Monitoring und metakognitiver Kontrolle erstellt, das theoriegeleitet und eng am Kodierschema der TTT entwickelt wurde. Im zweiten Auswertungsschritt werden die Logfile-Daten aktuell kodiert und ausgewertet. Die Leistung der Problemlöseaufgabe kann anhand des erstellten Schienenkreises operationalisiert werden.

Dabei werden folgende Hypothesen getestet: Kinder mit weniger Passungsversuchen (=höheres Monitoring und Kontrolle) (H1) weisen auch im Elternfragebogen höhere Metakognitionswerte auf (Konstruktvalidität), (H2) sind älter als Kinder mit vielen Passungsversuchen (Technik der bekannten Gruppen), (H3) erzielen besser Ergebnisse in der Problemlöseaufgabe (prädiktive Validität).

Zur Auswertung von H1 werden Korrelationen berechnet, um den Zusammenhang zwischen den beiden Metakognitionsmaßen zu bestimmen. Für die Technik der bekannten Gruppen wird mittels Regression untersucht, ob das Alter der Probanden einen Einfluss auf die mit den Logfiles gemessene Metakognition hat (H2). Zur Bestimmung der prädiktiven Validität wird anhand einer Regression berechnet, ob die so gemessene Metakognition die Leistung in der TTT vorhersagt (H3). Die Daten werden aktuell ausgewertet, so dass zum Zeitpunkt der Konferenz Ergebnisse zur Testung der drei Hypothesen vorliegen. Die Ergebnisse liefern neue Einsichten in innovative Verfahren zur Erfassung von Metakognition bei jungen Lernenden.

Macht authentisches Assessment einen Test schwerer? Das Beispiel von ECON 2022

Nina Welsandt, Hermann Josef Abs
Universität Duisburg-Essen

Theoretischer Hintergrund: Die Einführung des Unterrichtsfachs Wirtschaft-Politik ab dem Schuljahr 2020/21 in Nordrhein-Westfalen (vgl. MSB, 2021) ist ein Anlass für die Entwicklung eines neuen Wissenstests im Bereich der ökonomischen Bildung. Das Projekt „Ökonomische Bildung – Untersuchung zur Ausgangslage in Jahrgang 8“ (ECON 2022) bietet die Möglichkeit ein Authentic Assessment zu entwickeln.

Die zunehmende Digitalisierung führt vermehrt zum Einsatz von technologiebasierten Assessments in der ökonomischen Bildung. Auffällig ist dabei, dass der Fokus auf der Entwicklung von Authentic Assessments auch in jüngster Zeit noch nicht zentral ist (vgl. Welsandt & Abs, i.prep). Das zunehmende Potenzial des technologiebasierten Assessments (TBA), Bilder, Videos und Audiosequenzen als Teil des Assessments anzuzeigen, bietet jedoch neue Möglichkeiten, Testumgebungen authentisch zu gestalten (vgl. Janesick, 2006, S.4ff.; vgl. Jude & Wirth, 2007, S.49ff.; vgl. Koh, 2017). Der Hauptvorteil authentischer Testumgebungen besteht in dynamisch gestalteten Testaufgaben, die sich auf reale Handlungssituationen beziehen, deren Bewältigung alltagsrelevante Fähigkeiten erfordern (vgl. Janesick, 2006, S.4ff.).

Zu den Gründen, die gegen die Umsetzung von authentischen Testumgebungen angeführt werden, gehört neben dem hohen Aufwand in der Programmierung die Befürchtung, authentische Testaufgaben könnten eine höhere Informationsverarbeitungskapazität beanspruchen und so zusätzlich zur Schwierigkeit eines Tests beitragen (vgl. Aziz et al., 2020, S.763, vgl. Tanner, 2001, S.28).

Fragestellung:

Es stellt sich die Frage, wie die Authentizität einer Testumgebung mit möglichen schwierigkeitserzeugenden Merkmalen zusammenhängt und inwiefern die Testergebnisse von der Authentizität der Testumgebung beeinflusst sind?

Methode:

Für eine Expertenvalidierung des Tests nach Beck (2020) und Sangmeister et al. (2018) wurden die vier Aufgabenmerkmale (1) Domänenspezifität, (2) kognitivem Anspruchsniveau und (3) technischer Schwierigkeit der Itemmodellierung (vgl. Klotz et al., 2015; vgl. Winther, 2010) sowie (4) Authentizität operationalisiert. Zu jedem Testitem wurden Expertenratings (N=25) eingeholt. In einem ersten Schritt wurden die Einschätzungen der Aufgabenmerkmale (1) bis (3) mit der Einschätzung zur (4) Authentizität in Beziehung gesetzt. In einem zweiten Schritt wurden die vier Aufgabenmerkmale jeweils mit der im Rahmen eines Feldtests gemessenen Schwierigkeit der Aufgaben korreliert. Dabei interessiert, ob es gelungen ist, einen statistischen Zusammenhang zwischen der Authentizität und der Itemschwierigkeit zu vermeiden

Ergebnisse und ihre Bedeutung: Im Kontext der ökonomischen Bildung bilden technologiebasierte Assessments die Möglichkeit, typische Arbeits- und Denkabläufe realistisch zu konstruieren. Im Prozess der Konstruktion neuer Testinstrumente liegt der Fokus auf der Umsetzung einer möglichst authentischen und somit für die persönliche Lern- und Lebenswelt bedeutsamen Umgebung, die durch einen medialen Einsatz emotionale und motivationale Effekte haben kann.

Die Validierungsstudie zeigt, dass die Expertenratings der konzeptionell schwierigkeitserzeugenden Prinzipien (1) Domänenspezifität (.477**), (2) kognitives Anspruchsniveau (.445*) und (3) Itemmodellierung (.371*) signifikant mit der (4) Authentizität korrelieren. D.h. dass es gemäß des Urteils der Experten einen Zusammenhang zwischen den Schwierigkeit und Authentizität gibt. Weiterhin wurden die Experteneinschätzungen in Beziehung zur Itemschwierigkeit gesetzt, die mittels IRT-Skalierung auf der Basis von Feldtestdaten mit Schülerinnen und Schülern geschätzt wurde (vgl. Fortunati et al., i. prep). Von den drei schwierigkeitserzeugenden Prinzipien weist lediglich die Bewertung des (2) kognitiven Anspruchsniveaus (-.364*) eine signifikante Korrelation zur gemessenen Itemschwierigkeit auf. Es zeigt sich weiter, dass es keine signifikante Korrelation zwischen dem Aufgabenmerkmal der (4) Authentizität und der tatsächlichen Schwierigkeit einzelner Aufgaben im Feldtest gibt. Letzteres entspricht der Intention für die Umsetzung der authentischen Testumgebung.

Mit ECON 2022 wurde ein Assessment entwickelt, für das der Anspruch erhoben werden kann, das Kompetenzverständnis der ökonomischen Bildung in authentische Anforderungssituationen umgesetzt und verschiedene Facetten der ökonomischen Kompetenz abgebildet zu haben. Für die Diskussion und künftige Testentwicklungen ist die Methodik relevant, wie die Authentizität der Testumgebung bestimmt wurde und welche Verfahren geeignet sind, um den Einfluss von Authentizität auf Testschwierigkeit zu überprüfen.

Selbstkontrolliertes Verhalten von Kindern in einem computerbasierten simulierten Supermarkt

Philine Drake[1], Johannes Hartig[1], Manuel Froitzheim[2], Michael Schuhen[2], Gunnar Mau[3], Hanna Schramm-Klein[2]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany, [2]Universität Siegen, [3]Deutsche Hochschule für Gesundheit & Sport

Theoretischer Hintergrund

Die Anforderungen des Marktes mit ihrer Vielzahl von Reizen und Informationen können besonders für Kinder überwältigend sein, deren kognitive Fähigkeiten und Fertigkeiten noch nicht voll entwickelt

sind und denen es an Markterfahrung und -wissen fehlt (Mau et al., 2014). Um die Kaufentscheidungen von Kindern und die damit verbundenen Verhaltensweisen und Defizite besser zu verstehen, analysierten Mau et al. (2016) das Verhalten von Kindern in einer simulierten Supermarktumgebung. Sie zeigten, dass sich Kinder am Point of Sale oft anders verhalten, als sie es vorher beabsichtigt und erwartet hatten. Ihre Ergebnisse deuten darauf hin, dass Kinder Schwierigkeiten haben, eine Grundvoraussetzung für zielgerichtetes Konsumverhalten zu erfüllen, nämlich die notwendigen Handlungen umzusetzen, um ein bestimmtes Ziel zu erreichen.

Fragestellung

Obwohl die Bedeutung der Selbstkontrolle für die erfolgreiche Umsetzung von Konsumzielen hervorgehoben wurde (z. B. durch Vermeidung von Impulskäufen; Baumeister, 2002), gibt es noch keine Studie, die speziell die Vorgänge bei Kindern im Kaufprozess erfasst und das Ausmaß des selbstkontrollierten Verhaltens mit der erfolgreichen Umsetzung einer Kaufabsicht in Beziehung setzt.

Methode

Um diese Lücke zu schließen, haben wir eine computergestützte Supermarktsimulation verwendet, um das Einkaufsverhalten von Kindern am Point of Sale zu untersuchen. Bei dieser Aufgabe wurden die Kinder gebeten, über ein Tablet in der Supermarktsimulation einen Einkauf auf der Grundlage einer Einkaufsliste umzusetzen. Die Supermarktsimulation ist so konzipiert, dass auf der Verhaltensebene das Aufmerksamkeitsverhalten der Kinder aus den Logdaten der computerbasierten Aufgabe abgeleitet werden kann (Silberer, 2009). Das Aufmerksamkeitsverhalten umfasst die beobachtbare Aufmerksamkeit gegenüber Objekten in der Ladenumgebung, d.h. wie oft oder wie lange die Kinder einzelne Produkte betrachten. Die Supermarktsimulation wurde bewusst so gestaltet, dass sie Elemente enthält, die für die Leistungsaufgabe nicht erforderlich sind. Das Ausmaß, in dem sich die Kinder mit diesen irrelevanten Elementen beschäftigen, lässt sich an ihrem Aufmerksamkeitsverhalten ablesen und ermöglicht auf der Verhaltensebene eine Differenzierung zwischen Handlungen, die eher den Have-to-Zielen (gemäß der Aufgabe) oder den Want-to-Zielen im Sinne von Inzlicht et al. (2014) zuträglich sind.

Im Zentrum der Datenanalyse stand die Frage, ob die Kovarianz zwischen den Verhaltensindikatoren, von denen angenommen wird, dass sie Selbstkontrolle erfassen (z. B. das Ausmaß der Beschäftigung mit aufgabenirrelevanten Produkten), durch einen einzigen gemeinsamen Faktor erklärt werden kann und wie dieser Faktor mit dem Aufgabenerfolg, der Überwachung der Aufgabenerfüllung und den Ausgaben zusammenhängt. Eine Stichprobe von 117 Grundschulkindern erhielt eine Einkaufsliste und ein begrenztes Budget. Um Verhaltensindikatoren aus den Logdaten zu extrahieren, verwendeten wir den Finite-State-Machine-Ansatz (Kroehne & Goldhammer, 2018).

Ergebnisse und ihre Bedeutung

Es wurde eine eindimensionale konfirmatorische Faktorenanalyse (CFA) mit allen angenommenen Indikatoren analysiert. Das endgültige Modell für die Selbstkontrolle umfasste vier Variablen: Das zeitliche Ausmaß, mit dem die Kinder irrelevanten Regalen (S1) oder Produkten (S2) ihre Aufmerksamkeit schenkten, die Häufigkeit, mit der sie irrelevante Produkte kauften, die nicht auf der Einkaufsliste standen (S3) oder irrelevante Regale besuchten (S4). Das Modell zeigte eine

weitgehend gute Anpassung ($\chi 2(1) = 2.276$, $p = .131$, RMSEA = 0.105, 90% RMSEA CI [.28, .58], CFI = 0.993, TLI = 0.956, SRMR = .04). Lediglich der RMSEA überschritt das Cut-off-Kriterium. Der Aufgabenerfolg wurde anhand des Partial-Credit-Modells geschätzt. Die signifikante Korrelation zwischen dem Aufgabenerfolg (WLE) und dem Faktor für Selbstkontrolle ($r(113)=.44$, $p<.001$) deutet darauf hin, dass die Selbstkontrolle eine wichtige Rolle im Kaufprozess spielt. Unsere Ergebnisse zeigen auch, dass Kinder, die ihre Ausgaben (Ungenauigkeit der Schätzungen, $r(108)= -.35$, $p<.001$) und den Aufgabenerfolg ($r(111)=.40$, $p<.001$) sorgfältiger überwachten, tendenziell eine größere Selbstkontrolle bei der Aufgabenausführung zeigten. Unsere Untersuchung veranschaulicht, wie theoriebasierte Faktoren aus Logdaten computergestützter Aufgaben extrahiert werden können, und zeigt deren diagnostisches Potenzial auf, das zur Verbesserung der Qualität und Reichhaltigkeit psychologischer und bildungswissenschaftlicher Assessments genutzt werden kann.

E-15: Forschung und Praxis gemeinsam denken: Wie gelingt das Zusammenwirken von Bildungsforschung und Bildungspraxis?

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* R11 T00 D03

Forschung und Praxis gemeinsam denken: Wie gelingt das Zusammenwirken von Bildungsforschung und Bildungspraxis?

Chair(s): **Pia Sander** (Universität Duisburg-Essen, Germany), **Michael Kerres** (Universität Duisburg-Essen, Germany)
*Diskutant*in(nen):* **Josef Schrader** (Leibniz-Institut für Lebenslanges Lernen | Deutsches Institut für Erwachsenenbildung)

Das Bundesministerium für Bildung und Forschung (BMBF) hat seit einiger Zeit Metavorhaben eingeführt, um Projekte von Förderlinien im Hinblick auf Fragen der Zusammenarbeit mit der Praxis in den Phasen eines Vorhabens, der Wissenschaftskommunikation und des Transfers zu unterstützen (vgl. z.B. Diekmann, 2022). Auch anderen Fördereinrichtungen wird es immer wichtiger, dass diese Thematik in den Vorhaben (oder durch Begleitvorhaben) explizit angesprochen wird (Spoden & Schrader, 2021). Insbesondere in der Bildung besteht jedoch weiterhin die Frage, wie Bildungspraxis und Bildungsforschung zusammenwirken können, um Ziele beider Akteure einzulösen (Gräsel et al., 2019). Eine mögliche Erklärung dafür liegt in den unterschiedlichen Logiken der beiden Systeme: So beantwortet Bildungsforschung Fragen des Wissenschaftssystems, während Bildungspraxis gesellschaftliche Anliegen des Bildungssystems adressiert (Luhmann & Schorr, 1982). Gestaltungsorientierte Bildungsforschung, aber auch andere Ansätze der Forschung-Praxis-Kooperationen und des Transfers bieten Ansatzpunkte, wie Forschung im Zusammenwirken mit Praxis gestaltet werden kann (Kerres, Sander & Waffner, 2022; Schrader et al., 2020). Solche Ansätze haben zum Ziel, dass Forschung nicht allein Fragen der Wissenschaft beantwortet, sondern auch einen Nutzen für Praxis bietet, indem sie Lösungen für konkrete Bildungsanliegen erarbeitet (Tulodziecki et al., 2018).

In dem Symposium wird der Frage nachgegangen, wie Forschung und Praxis als ko-konstruktiver Prozess gedacht werden kann. In vier Beiträgen werden Ergebnisse aus empirischen Studien vorgestellt, die Methoden und Wege des erfolgreichen Zusammenwirkens von Bildungsforschung und Bildungspraxis aufzeigen und Erfahrungen reflektieren.

Beitrag 1 untersucht basierend auf dem Rahmenmodell von Farley-Ripple et al. (2018), welche Transferaktivitäten Forschende in verschiedenen Phasen von empirischen Projekten zur Bildungsforschung durchführen. Die Ergebnisse der Interviewstudie werden in Projektprofilen verdeutlicht.

Beitrag 2 widmet sich der Frage, wie innovative Qualifizierungsangebote in der beruflichen Bildung in Wissenschafts-Praxis-Kooperationen entwickelt werden können. Hierzu werden verschiedenen qualitative Erhebungen in Projekten der BMBF-Förderlinie InnoVET ausgewertet und Erkenntnisse zu Innovationstransfer und Translationsprozessen erlangt.

Beitrag 3 präsentiert die Ergebnisse einer Längsschnittbefragung des BMBF-Metavorhabens zur Digitalisierung im Bildungsbereich aus 49 Förderprojekten zum Zusammenwirken von Forschung und Praxis. In drei Wellen wurden Forschende zu ihren geplanten und durchgeführten Formen des Dialogs mit der Praxis sowie den erzielten Effekten des Austauschs befragt.

Beitrag 4 nimmt schließlich die Praxisperspektive genauer in den Blick und erforscht die Gelingensbedingungen für ein erfolgreiches Zusammenwirken von Bildungsforschung und -Praxis aus

Perspektive von *Citizen Science*. Hierfür werden die Daten aus sieben Interviews mit in Forschungsprojekten beteiligte Bürger:innen vorgestellt und diskutiert.

In der anschließenden Diskussion wird die Frage weiter erörtert, wie durch einen ko-konstruktiven Prozess Herausforderungen in Bildungskontexten bearbeitet werden können und wo Limitationen bestehen.

Beiträge des Symposiums

Transferaktivitäten in Projekten der empirischen Bildungsforschung: Ergebnisse einer Interviewstudie und Entwicklung eines theoriegeleiteten Befragungsinstruments

Ulrike Hartmann, Mareike Kunter
DIPF - Leibniz-Institut für Bildungsforschung und Bildungsinformation

Dass Bildungsforschung praktisch relevante Erkenntnisse hervorbringen soll, die sich in Veränderungen in der schulischen Praxis niederschlagen, ist eine so alte wie vieldiskutierte Forderung. Begriffe wie Theorie-Praxis-Problem oder Research Practice Gap verdeutlichen allerdings, dass mit dieser Forderung vielfältige Herausforderungen verbunden sind.

Forschungsbefunde thematisieren hauptsächlich die Nutzung von bildungswissenschaftlichen Erkenntnissen. Diese zeigen recht einheitlich, dass Lehrkräfte im Berufsalltag nur gelegentlich wissenschaftliche Befunde rezipieren und sich hauptsächlich auf ihre eigenen Erfahrungen oder Ratschläge anderer Kolleg*innen verlassen (Cain, 2016; Hetmanek et al., 2015; Hinzke, Gesang & Besa, 2020; Joram, Gabriele & Walton, 2020; Lysenko et al., 2015). Zu den Herausforderungen aufseiten der Bildungsforschung überwiegen bislang theoretische Reflexionen darüber, welche Erkenntnisse Bildungsforschung hervorbringt und inwiefern diese relevant für die pädagogische Praxis sind (z.B. Prenzel, 2010; Renkl, 2022; Stark, 2017; Wilkes & Stark, 2022).

Die hier vorgestellte Untersuchung thematisiert die Angebots-Seite und zeigt, welche Aktivitäten in Projekten der empirischen Bildungsforschung realisiert werden, um einen Transfer im Hinblick auf die Bildungspraxis zu ermöglichen. Theoretisch wird auf das Rahmenmodell von Farley-Ripple und Kolleg*innen (2018) rekurriert. Dieses macht Annahmen darüber, wie verschiedene Dimensionen von Transferaktivitäten mit dem Ausmaß der Research Practice Gap zusammenhängen.

Wir gehen der übergeordneten Frage nach, welche Formen von Transferaktivitäten von Forschenden in Projekten der Bildungsforschung realisiert werden. Es wird untersucht, welche Ausprägungen von Transferdimensionen nach dem Rahmenmodell von Farley-Ripple et al. (2018) sich identifizieren lassen, und inwieweit sich ausgehend von diesen Transferdimensionen Profiltypen bestimmen lassen.

Methode

Die hier vorgestellten Ergebnisse beziehen sich auf Eingangsinterviews eines längerfristig angelegten Entwicklungsprojektes, an dem sieben Leitungspersonen (Post-Docs und Professor*innen) aus Bildungsforschungsprojekten teilnahmen. Alle Projekte sind drittmittelfinanziert und wurden im Rahmen kompetitiver Ausschreibungsverfahren eingeworben (DFG (3x), Bundesministerien (2x), EU-Förderung (1x), Sonstige (1x). Die Projekte arbeiten vorwiegend im quantitativen Methodenspektrum (z.B. experimentelle Studien, Large Scale Assessments, Befragungsstudien), wobei mehrere Projekte zusätzlich Datenquellen wie Interviews oder Gruppendiskussionen einbeziehen. Die

leitfadengestützten Interviews dauerten ca. 60 Minuten. Zu vier Dimensionen aus dem Rahmenmodell (Dissemination, Partizipation, Interpretation, Häufigkeit) wurden Leitfragen gestellt und durch konkrete Rückfragen ergänzt. Interviews wurden transkribiert und qualitativ-quantifizierend anhand eines theoriegeleiteten Kodierschemas ausgewertet (Hartmann & Kunter, 2022). Dabei wurden 30% des Materials von zwei studentischen Mitarbeiterinnen kodiert. Für alle Dimensionen und zugehörige Kategorien wurden gute bis sehr gute Übereinstimmungen erzielt (0.81-1.00). Um Profile der Projekte anhand der Dimensionen des Praxistransfers erstellen zu können, wurden die Einzelkodierungen jedes Projektes zu einem Gesamtwert pro Kategorie aggregiert. Jede Kategorie wurde dafür in drei Niveaustufen eingeteilt (0 = gering, 1 = mittel, 2 = hoch), um eine vergleichende Profildarstellung über Kategorien hinweg zu ermöglichen.

Ergebnisse und Implikationen

Für die Dimension Dissemination zeigt sich, dass vorwiegend unidirektionale Vermittlungsstrategien gewählt werden und kaum interaktive Transferformate vorkommen. Die Dimension Partizipation wird in den Projekten hauptsächlich durch individuelle Brokering-Aktivitäten aufgegriffen, und deutlich seltener durch eine systematisch organisierte Mitarbeit von Praxispersonen im Projektteam. Bezüglich der Dimension Interpretation ist ein hohes Reflexionspotenzial der befragten Projektleitungen erkennbar, was Voraussetzungen und Verständnisschwierigkeiten von Praxiszielgruppen betrifft – dies schlägt sich jedoch vergleichsweise selten in elaborierten Vermittlungsstrategien nieder. Die Ergebnisse zur Dimension Häufigkeit zeigen, dass Transferaktivitäten in der Mehrzahl der Projekte lediglich in einzelnen Projektphasen verfolgt werden. Profilanalysen verdeutlichen den Mehrwert einer dimensionalen Betrachtung von Praxisbezügen. Es lassen sich in unserer Stichprobe keine Projekte finden, die in ihren Praxisbezügen als durchgängig hoch oder niedrig eingeschätzt werden. Vielmehr findet sich eine große Varianz sowohl zwischen den Projekten als auch innerhalb jedes einzelnen Projektes.

Im Vortrag werden Anlage und Ergebnisse der Interviewstudie dargestellt sowie die darauf basierende Entwicklung eines Fragebogeninstrumentes, das Ausprägungen von und Einstellungen zu Transferaktivitäten nach dem Rahmenmodell von Farley-Ripple et al. ökonomisch erfassbar macht und an größeren Stichproben von Bildungsforschungsprojekten eingesetzt werden kann, um die Befunde der Interviewstudie zu validieren und vergleichend zu analysieren.

Integrative Vernetzung von Forschung und Praxis zur Unterstützung von Innovationen in der Berufsbildung

Hubert Ertl, **Nina Peitz**
Bundesinstitut für Berufsbildung (BIBB)

In Bezug auf die Frage nach dem Nutzen der Bildungsforschung und im Besonderen danach, wie bildungspolitische Förderprogramme ausgerichtet sein müssen (Kerres et al., 2022), soll der Transfer von Neuem über diese Programme hinausgedacht werden (BIBB, 2021) und in Metaprojekten selbst zum Forschungsgegenstand werden (Ehrke, 2017). Die Transferproblematik von Wissen und Ergebnissen in der Bildungsforschung und -praxis gewinnt in Bildungsvorhaben immer mehr an Bedeutung (Stockmann, 2021; Gräsel et al., 2005). Statt Transferaktivitäten am Ende von Projekten zu realisieren und evaluieren, wird zunehmend gefordert, den Transfer von Beginn an zu gestalten

und zu erproben, sowie Akteure oberhalb der Projektebene in den Prozess einzubinden (Eckelt, 2018). Daher bedarf es neuer Unterstützungsmechanismen, die sich dieser Herausforderung gerade auch in komplexen Programmen annehmen (Ertl, 2022; Stockmann, 2021). Im Rahmen des durch das Bundesministerium für Bildung und Forschung initiierten und durch das Bundesinstitut für Berufsbildung durchgeführten Programms InnoVET (Innovationswettbewerb für eine exzellente Berufliche Bildung) untersucht ein Begleitforschungsverbund übergeordnet Fragen zu Innovationstransfer und Translationsprozessen.

Eine Zielsetzung des InnoVET-Programms ist es, mittels neuer Fortbildungsabschlüsse durch die Vernetzung der Bildungsbereiche (besonders der beruflichen und hochschulischen Bildung) und die Kooperationen zwischen Lernorten letztlich die Durchlässigkeit im Berufsbildungssystem zu verbessern. Gesellschaftliche Diskurse verdeutlichen, dass mit Durchlässigkeit neben der formalen Ermöglichung von Zu- und Übergängen (Weber, 2012) zwischen den Bildungsbereichen ebenso Ansprüche wie Modernisierung und Innovationen zu individuell flexiblen Lernwegen für Lernende verbunden sind (Hemkes & Wilbers, 2019). Der Deutsche Qualifikationsrahmen (DQR) besteht aus acht Kompetenzniveaus, die sich den Bildungsqualifikationen in Deutschland zuordnen lassen. Die Niveau-Stufe 5 des DQR spielt eine große Rolle beim Thema Durchlässigkeit, da sie eine Brücke zwischen beruflichen und hochschulischen Bildungsgängen und -abschlüssen herstellen kann (Wilbers, 2022). Die Stufen 4 (mindestens 3-jährige Ausbildungsberufe) und 6 (Bachelor-Niveau) werden miteinander verbunden. Ein wichtiger Ansatz hierfür ist die einheitliche Fortbildungsabschlussbezeichnung „Geprüfte/r Berufsspezialist/in“ mit Bezug zur DQR-Stufe 5, durch die sich ein neuer Weg zur beruflichen Weiterentwicklung eröffnet.

Welche Ergebnisse zur Verbesserung von Durchlässigkeit zwischen Bildungsbereichen sind von wem mit welchen Interessen, Instrumenten und Zielsetzungen wohin transferierbar?

• Welche innovativen Fortbildungen auf der DQR-Stufe 5 (Berufsspezialisten) entstehen durch welche Arten der Kooperation (Bildungsforschung und -praxis) in ausgewählten InnoVET-Projekten?

• Wie werden sie in den Projekten systemisch integriert? Welche Wirkung entfalten sie?

• Was sind Gelingensbedingungen (z.B. förderliche rechtliche Rahmenbedingungen) bei der Entwicklung und Durchführung von Programmen, die auf die Etablierung von neuen Fortbildungen auf DQR-Stufe 5 als Treiber für Durchlässigkeit des Berufsbildungssystems abzielen?

InnoVET weist mit seinem großflächig angelegten Netzwerk mit über achtzig Verbundpartnern eine besondere Förderstruktur auf (BIBB, 2021). Aufgrund dieser sowie der Zielsetzung der Entwicklung und Verbreitung von Innovationen eignet sich für die übergeordnete Begleitforschung eine gestaltungsorientierte Vorgehensweise, die von der Zusammenarbeit zwischen Wissenschaft und Praxis zur Lösung eines offenen Problems der Bildungspraxis geprägt ist, und gleichermaßen einen praktischen wie theoretischen Output liefert (Cobb et al., 2003, Anderson & Shattuck, 2012).

Mittels Fallstudien zu InnoVET-Projekten (z. B. BIRD) wird entlang der o.g. Fragen untersucht, wie die innovativen Qualifizierungsangebote auf der DQR-Stufe 5 durch eine Wissenschafts-Praxis-Kooperation entwickelt und in der Praxis erprobt werden. Nach einer Dokumentenanalyse der Projektanträge, Berichte und wissenschaftlichen Papiere (vgl. Wilbers, 2022) erfolgen eine Reihe von

qualitativen Erhebungen. Es werden leitfadengestützte Interviews mit Projektbeteiligten geführt, deren Ergebnisse in einem anschließenden Workshop mit Experten und Projektbeteiligten aus Wissenschaft und Praxis vertieft werden. Die Analyse der Daten erfolgt mittels qualitativer Inhaltsanalyse zur Beantwortung der Forschungsfrage(n), erste Ergebnisse daraus werden präsentiert.

Die Ergebnisse zeigen im Symposium anhand von InnoVET-Projektbeispielen auf, wie durch Maßnahmen zur Entwicklung und Etablierung der Qualifizierung zum Berufsspezialisten die Durchlässigkeit im Berufsbildungssystem verbessert werden kann. Insbesondere Schlussfolgerungen zu den Interessen und Instrumenten zum Transfer der Neuerungen im Gefüge der Wissenschafts-Politik-Praxis-Kooperation sind wichtige Indizien, wie Förderprogramme zukünftig gestaltet werden sollen.

Kommunikation und Transfer von Forschenden mit der Praxis in Projekten zur Bildung in der digitalen Welt

Katja Buntins, Pia Sander
Universität Duisburg-Essen, Germany

In einem schnelllebigen Forschungsfeld wie der Mediendidaktik, braucht es eine funktionierende Kommunikation zwischen Praktiker*innen und Forschenden. Wissenschaftskommunikation sowie Wissenstransfer zwischen Bildungsforschung und Bildungspraxis gewinnen daher zunehmend an Bedeutung (Weingart et. al, 2019). Dieses Desiderat adressiert die Politik in ihren Ausschreibungen, indem sie neben Forschungsprojekten zusätzliche Metavorhaben fördert, um die Zusammenarbeit zwischen Forschung und Praxis zu unterstützen. Solch ein zielgerichtetes Zusammenwirken zwischen unterschiedlichen Akteuren soll eine Kultur der Forschung-Praxis-Kooperation positiv beeinflussen (Schrader et al., 2020).

Erkenntnisse zu der Wirkung von Wissenschaftskommunikation in der deutschen Mediendidaktik liefert das vom Bundesministerium für Bildungsforschung (BMBF) geförderte Metavorhaben Digitalisierung im Bildungsbereich (digi-EBF). Dies vernetzt Wissenschaftler*innen in den Förderlinien zu Digitalisierung untereinander und bringt sie in den Dialog mit Bildungspraxis (Diekmann, 2022). Darüber hinaus beforscht es die Anlage einer gestaltungsorientierten Bildungsforschung. Diese geht davon aus, dass ein erfolgreiches Zusammenwirken zwischen Bildungsforschung und -praxis durch die Kommunikation der jeweiligen Akteure geformt wird (Kerres, Sander & Waffner, 2022). Ziel dieser Kommunikation muss es sein, den Nutzen des Forschungsvorhaben für beide Seiten zu klären und über die Projektphasen hinweg aufrechtzuerhalten. Nur so können die differenzierten Handlungslogiken (siehe auch Luhmann & Schorr, 1982) von Bildungsforschung und –praxis aufgelöst werden und ein ko-konstruktiver Prozess entstehen.

Der Beitrag betrachtet die Frage, ob und inwiefern die geförderten Projekte in digi-EBF mit der Praxis kommunizieren und welche Transferaktivitäten sie durchführen. Es wird aufgezeigt wie und mit wem kommuniziert wird sowie welche Prädiktoren für Kommunikation festzustellen sind. Hierbei werden auch längsschnittliche Entwicklungen berücksichtigt.

Methode

Die in digi-EBF beteiligten Projekte wurden in drei Wellen zu ihren Transferaktivitäten befragt: am Anfang, zur Mitte und am Ende ihrer Laufzeit. Es wurden alle wissenschaftlichen Mitarbeitenden und

Projektleitungen eingeladen an den Umfragen teilzunehmen. Die zugrunde liegende mulitlevel cross-classified Struktur (Goldstein, 1994) wird auf Grund der geringen Stichprobengröße innerhalb eines Projektteils mit drei verschiedenen Datensätzen beschrieben. Innerhalb der drei Wellen wurden insgesamt 49 Projekte mit 103 Teilprojekten zu mindestens einem Zeitpunkt befragt. Auf Ebene der Statusgruppen wurden 29 Fälle auf Ebene der wissenschaftlichen Mitarbeitenden und 50 Fälle auf der Ebene der Projektleitungen berücksichtigt.

In allen Wellen wurden die Projekte dazu befragt, welche Kommunikationswege mit der Praxis sie planen (Welle 1) bzw. verwendet haben (Welle 2 und 3). In Welle 2 und 3 wurde darüber hinaus erfasst, mit welchen Praktiker*innen kommuniziert wurde und was der wahrgenommene Nutzen dieser Kommunikation war. In Welle 3 wurde außerdem nach der Einstellung zur Praxiskommunikation gefragt.

Ergebnisse & Implikationen

Die Ergebnisse zeigen, dass am Ende eines Projektes jedes Teilprojekt mindestens ein Kommunikationsformat mit der Praxis gewählt hat. Hierbei wurden kurzweiligere Formate, wie eine Datenerhebung in der Praxis, Vorträge und informeller Austausch häufiger gewählt. Jedoch wurden auch bei 40,7 Prozent der Projektteile eine gemeinsame Bearbeitung vorgenommen und bei 43,2 Prozent die erarbeiteten Ergebnisse mit der Praxis validiert. Die Intensivität der Kommunikation bleibt zwischen der zweiten und der dritten Welle im Längsschnitt konstant, jedoch nimmt die Häufigkeit der angebenden Aktivitäten im Vergleich zwischen den Wellen zu. Die Häufigkeit und Intensivität des Austausches hängt mit dem Nutzen und der Einstellung zur Praxiskommunikation zusammen. Auch gibt es einen Zusammenhang mit der Forschungsmethode.

Kommunikation als zentrales Element einer gestaltungsorientierten Bildungsforschung ist, insbesondere in einem schnell wachsenden Feld wie Bildung in einer digitalen Welt, essentiell. Die präsentierten Ergebnisse zeigen, dass die Formate und Intensität von Kommunikation in den Projekten in digi-EBF durchaus stattfinden. Vor allem dann, wenn transferaffine Wissenschaftlerinnen involviert sind. Jedoch gibt es Hinweise darauf, dass teils ein veraltetes Verständnis von Kommunikation vorherrschend ist, in dem Wissenschaft linear und paternalistisch Richtung Praxis kommuniziert (Buchini & Trench, 2014). Um zu evaluieren, inwieweit wissenschaftliche Erkenntnisse tatsächlich von der Praxis aufgenommen und verarbeitet werden, ist die Erforschung der Wahrnehmung und Handlungsweisen von Praktiker*innen notwendig.

Citizen Science als Forschungsansatz für erfolgreiches Zusammenwirken von Forschung und Praxis

Lisa Birnbaum, Stephan Kröner
Friedrich-Alexander-Universität

Wissenschaftskommunikation und Praxistransfer gewinnen immer stärker an Bedeutung. Citizen Science ermöglicht es, diese über den gesamten Forschungsprozess hinweg zu fördern. Unter Citizen Science versteht man das gemeinsame Forschen von hauptberuflich an Forschungseinrichtungen tätigen Wissenschaftlerinnen und Wissenschaftlern und interessierten Bürgerinnen und Bürgern (Haklay et al., 2021). Bislang sind die meisten Citizen-Science-Projekte naturwissenschaftlich ausgerichtet (Kullenberg & Kasperowski, 2016). Außerdem beschränkt sich die

Beteiligung von Citizen Scientists überwiegend auf Hilfsarbeiten im Rahmen der Datenerhebung (z. B. Probenentnahme, Beobachtungen). Wünschenswert ist zum einen die verstärkte Ausweitung auf Sozial- und Geisteswissenschaften, zum anderen eine empowernde, sogenannte ko-kreative Einbindung der Bürgerinnen und Bürger auf Augenhöhe mit den hauptberuflich Forschenden über den gesamten Forschungsprozess hinweg, d.h. von der Entwicklung der Forschungsfrage über Datenerhebung und Datenanalyse bis hin zur Ergebnispräsentation (Robinson et al., 2021). Gerade in der Bildungsforschung bieten Citizen-Science-Projekte die Möglichkeit, Akteuren der Bildungspraxis Mitwirkung über den gesamten Forschungsprozess hinweg zu ermöglichen. Diese kann als informeller Bildungsprozess verstanden werden (Edwards, 2015; Eitzel et al., 2017).

Welche Erfahrungen machen Citizen Scientists in ko-kreativen Forschungsprojekten und welche Implikationen ergeben sich daraus für das erfolgreiche Zusammenwirken von Forschung und Praxis im Bildungsbereich? Diese Frage wird anhand von Interviewdaten mit Akteuren aus einem ko-kreativen Citizen-Science-Projekt bearbeitet. In diesem Projekt haben interessierte Bürgerinnen und Bürger gemeinsam mit Forschenden untersucht, inwiefern eine Lokalzeitung in einfacher Sprache Menschen mit niedrigem Leseselbstkonzept als Leserschaft erreicht und deren wahrgenommene gesellschaftliche Teilhabe stärken kann – sowie inwieweit andere Zielgruppen die Nachrichten auch in einfacher Sprache weiterhin lesen. Im Rahmen des Projekts waren die interviewten Bürgerinnen und Bürger redaktionell und als Autorinnen und Autoren an der Erstellung einer Lokalzeitung in einfacher Sprache beteiligt. Des Weiteren haben sie in dem zugehörigen Forschungsprozess auf Augenhöhe an der Entwicklung, Auswertung und Interpretation eines darauf bezogenen Fragebogens mitgewirkt. Die darauf bezogenen Projektergebnisse wurden abschließend der Öffentlichkeit präsentiert. Alle diese Schritte erfolgten ko-kreativ in Zusammenarbeit von Bürgerinnen und Bürger gemeinsam mit akademischen Forschenden.

Basierend auf n = 7 Interviews mit Citizen Scientists sowie informeller Beobachtungen der am Forschungsprozess beteiligten hauptberuflich Forschenden wird exploriert, was die Citizen Scientists des untersuchten Projekts motiviert hat, sich in das Forschungsprojekt einzubringen und was sie Gelingensbedingungen für erfolgreiches Zusammenwirken von Bildungsforschung und Praxis sind. Die qualitativ inhaltsanalytische Auswertung der Interviews zeigte, dass die Citizen Scientists durch das Projekt viel gelernt haben, sowohl zum Thema einfache Sprache als auch über empirische Bildungsforschung. Die Citizen Scientists hatten Gelegenheit, sich aktiv an empirischer Bildungsforschung zu beteiligen und sich mit dem noch relativ neuen, jedoch gesellschaftlich äußerst relevanten Thema einfache Sprache auseinanderzusetzen. Sie haben den beiden beteiligten hauptberuflich Forschenden zufolge im Rahmen der Fragebogenentwicklung mit ihrer Feldkenntnis substantiell dazu beigetragen, den zu entwickelnden Fragebogen optimal auf die Zielgruppe zuzuschneiden. Zudem wurde durch den Citizen-Science-Ansatz kontinuierlich über das Projekt in der Öffentlichkeit kommuniziert, so dass Lehrkräfte auf das Thema aufmerksam wurden, die daraufhin Fortbildungen angefragt haben. Dies verweist darauf, dass Citizen-Science-Projekte das Interesse an Themen der Bildungsforschung in der Zivilgesellschaft stärken können. Sie können so zum Ausgangspunkt für nachhaltige und durch wechselseitiges Vertrauen geprägte Wissenschafts-Praxis-Kooperationen werden. Eine aktive Zusammenarbeit auf Augenhöhe zwischen Citizen Scientists und

hauptberuflichen Wissenschaftlerinnen und Wissenschaftlern scheint wesentlich zu sein für den Erfolg von Citizen Science als Ansatz für gelingendes Zusammenwirken von Wissenschaft und Praxis. Es wird diskutiert, inwiefern die Erkenntnisse aus dem vorliegenden Projekt auf weitere Citizen-Science-Projekte übertragbar sind und welches Potenzial Citizen Science für das Zusammenwirken von Bildungsforschung und Bildungspraxis aufweist.

E-16: (Virtuelles) kooperatives Arbeiten und Lernen

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* T03 R03 D89

(Virtuelles) kooperatives Arbeiten und Lernen

Chair(s): **Christin Siegfried** (Georg-August-Universität Göttingen, Germany), **Henrik Bellhäuser** (Johannes Gutenberg-Universität Mainz)
*Diskutant*in(nen):* **Esther Winther** (Universität Duisburg-Essen)

Kooperatives Arbeiten und Lernen findet mit der Erwartung statt, dass Menschen mehr zusammen erreichen als alleine (Johnson et al., 2000). Entscheidend hierfür ist die geeignete Zusammensetzung der Teams (Bellhäuser et al., 2018) und die Partizipation jedes Teammitglieds durch Kommunikation (schriftlich, mündlich, non-verbal), indem Ideen, Erfahrungen und Wissen möglichst divers und elaboriert mit den anderen geteilt wird (Siegfried & Heinrichs, 2020; Siegfried, 2021). Allerdings wird Kooperation durch das World Wide Web und andere Technologien erweitert und substantiell verändert. So ermöglicht die stets verfügbare Verbindung zum Internet immer und überall in Kontakt mit Peers zu sein (Grieve et al., 2013). Auch die asynchrone Kooperation, welche die die Reflexion (eigener) Beiträge begünstigen kann (Robinson & Hullinger, 2008), wird durch den digitalen Kontext erleichtert. Demgegenüber kann die Konfrontation mit verschiedenen Kommunikations- und Informationsströmen zu kognitiver Belastung führen (Picciano, 2002; Leutner, et al., 2014). Technische Störung oder eigenes (Des-)Interesse macht es im virtuellen Raum leicht, Teamgesprächen fernzubleiben oder sie zu verlassen, was Frustration (Downing et al., 2007) auslösen kann. Schriftlicher Austausch kann zudem Missverständnisse begünstigen, wenn z.B. detaillierte Begründungen fehlen (Gibson & Gibbs 2006). Längeren Videokonferenzen können zudem ein Gefühl der Erschöpfung, auch als Zoom-Fatigue bezeichnet, auslösen (Nesher et al., 2021)

Das Symposium widmet sich förderlichen und hinderliche Faktoren im Rahmen virtuellen kooperativen Arbeitens und Lernens. Vor diesem gemeinsamen Hintergrund werden durch drei Beiträge in verschiedenen Sachkontexten folgende Fragen in den Blick genommen: 1. Welche Einflussfaktoren können für eine erfolgreiche virtuelle Teamarbeiten im Arbeitskontext identifiziert werden, 2. Inwiefern unterscheiden sich Organisationsprozesse von Gruppenarbeiten im Schulkontext in Präsenz von solchen in hybriden, online Formaten, und 3. Inwiefern können algorithmische Gruppenformation als Alternative zu selbstgebildeten Lerngruppen betrachtet werden. Die Beiträge bedienen sich dabei in interdisziplinärer Weise unterschiedlicher methodischer Herangehensweisen, die einen breiteren Blick auf den Forschungsgegenstand erlaubt.

Der erste Beitrag nutzt eine systematische Literaturrecherche, um Einflussfaktoren auf eine gelingende virtuelle und hybride Teamarbeit sowohl auf Input- und Prozessebene zu identifizieren. Hierfür wird das 3P Modell herangezogen.

Der zweite Beitrag analysiert mithilfe von Videotranskripten, wie die Gruppenorganisationsprozesse in Gruppen in Präsenz, hybrid und online ablaufen. In einem zweiten Schritt findet eine Gegenüberstellung der Organisationsprozesse statt, um Unterschiede entsprechend des Settings zu identifizieren.

Im dritten Beitrag wird untersucht, welche Einstellung Studierende gegenüber einem Algorithmus haben, der kriterienbasiert die Gruppenformation für sie übernimmt und inwiefern sie sich in einer realen Lehrveranstaltung für diese Option entscheiden.

Die abschließende Diskussion führt die gewonnenen Erkenntnisse zusammen und gibt einen Ausblick auf zukünftige Forschungsdesiderata.

Beiträge des Symposiums

Same same, but different – What it takes to successfully collaborate in digital environments

Ascha Ahmed, Viola Deutscher
University of Mannheim

Theoretischer Hintergrund

Aktuelle Studien lassen annehmen, dass viele Interaktionen am Arbeitsplatz auch nach COVID-19 weiterhin virtuell oder hybriderfolgen werden (McKinsey, 2020; Gartner, 2020). Virtuelle oder hybride Zusammenarbeit am Arbeitsplatz findet in Teams statt, die sich über die geografische Verteilung ihrer Mitglieder sowie über den Einsatz von Kommunikationstechnologien definiert, synchron oder asynchron auf ein gemeinsames Ziel hinzuarbeiten (Morrison Smith & Ruiz, 2020). Als theoretische Grundlage für virtuelle Zusammenarbeit lässt sich eine zusammengeführte Konzeptualisierung von Afferbach (2020), Gratton (2021) und Sangwan et al. (2020) entlang der folgenden Dimensionen heranziehen:

• Physische Distanz

• Zeitliche Distanz (ursprünglich „cognitive distance“)

• Bedarf an Kommunikationstechnologie

Unter Berücksichtigung dieser Faktoren wurden drei verschiedene Typen von Kollaboration definiert: traditionelle (face-to-face), hybride und rein virtuelle Kollaboration.

Die neue Art des Arbeitens ermöglicht es Organisationen, die besten Talente rund um den Globus zu gewinnen und zu vernetzen, unabhängig davon, wo sie sich befinden. Remote-Arbeiten haben sich nicht nur als kosteneffizient, agil und umweltfreundlich erwiesen, sondern bringt auch Flexibilität mit sich, die von Arbeitgebern und ArbeitnehmerInnen in der Industrie weiterhin gefragt sein wird (WEF, 2020; Wanda Curlee & PMP, 2011; Lund et al., 2021). Trotz der Vorteile hat die Forschung in der Vergangenheit bereits gezeigt, dass räumlich verstreute und zeitlich flexible Teams Schwächen in Bezug auf effektiven Wissensaustausch, Koordination, Innovation, Kreativität, Problemlösung, Vertrauensbildung, Teamdynamik und Entscheidungsfindung aufweisen können (Cramton, 2001; Cramton & Webber, 2004; Hill et al., 2009; Hinds & Mortensen, 2005; Joshi et al., 2009; O‘Leary & Cummings, 2007; Hill & Bartol, 2016; Thompson, 2021). Angesichts dieser Schwächen bedarf es konzeptionell an Qualitätsmodellen zu erfolgreicher Kollaboration bzw. Teamperformance und an empirischen Erkenntnissen zu den Erfolgsfaktoren. Den Rahmen hierfür können sog. 3P-Modelle bilden (z.B. Biggs, 1999; Tynjälä, 2013), welche sich mit gegebenen Ausgangsbedingungen, sowie relevanten Prozessfaktoren und deren Einfluss auf bestimmte Lern- und Arbeitsergebnisse beschäftigen (Input-Process-Outcome). Der eingereichte Beitrag beschäftigt sich im Rahmen eines solchen 3P Modells mit den Einflüssen für eine erfolgreiche Teamperformance unter Berücksichtigung der Besonderheiten von virtuellen oder hybriden Teams.

Fragestellung

Der vorliegende Beitrag untersucht folgende Fragestellungen:

RQ 1: Was bedeutet erfolgreiche virtuelle und hybride Zusammenarbeit am Arbeitsplatz bzw. welche Outcome-Dimensionen kennzeichnen sie?

RQ 2: Welche Faktoren auf Input- und Prozessebene bestimmen den Erfolg von virtueller und hybrider Zusammenarbeit?

Methode

Zur Beantwortung der oben genannten Forschungsfragen wurde eine systematische Literaturrecherche nach den PRISMA-Richtlinien durchgeführt (Liberati et al., 2009). Die Literatursammlung erfolgte im Mai-Juli 2022 unter Verwendung der Datenbanken ProQuest, Taylor & Francis Online -. Journals, EBSCOhost, Scirus, SpringerLink, SAGE Journals und Emerald Management eJournals. Die Suchbegriffe wurden aus den Forschungsfragen abgeleitet und im Rahmen einer Peer-Diskussion erweitert. Nach Reduzierung der Ergebnisse auf peer-reviewed Zeitschriften und eBooks, wurden Auswahlkriterien wie „virtuelles“ oder „hybrides Setting“ und „Arbeitsplatzumgebung“ verwendet, um die Relevanz der ausgewählten Literatur zu gewährleisten.

Ergebnisse und ihre Bedeutung

Nach der Sichtung von 448 Studien wurden 126 in die qualitative Synthese einbezogen. Auf Basis dieser Ergebnisse wird ein angepasstes bzw. erweitertes Qualitätsmodell für erfolgreiche Kollaboration/Teamperformance am virtuellen oder hybriden Arbeitsplatz vorgeschlagen. Während die Input-Faktoren in bereits bestehenden Qualitätsmodellen in individual-level und group-level differenziert werden, wurde in unserem Vorschlag eine weitere Ebene ergänzt: leadership-level. Dies wird gerechtfertigt mit der zunehmenden abgegrenzten Betrachtung des Führungsstils und dessen Bedeutung für den Prozess der Kollaboration selbst. So wird aufgrund der räumlichen Dispersion und der damit einhergehenden sinkenden Kontrolle durch die Führungskraft ein transformationeller oder befähigender Führungsstil empfohlen, der das Team während des Kollaborationsprozesses ermächtigt eigenständig und agil zu handeln. Aufgrund einer fehlenden Berücksichtigung der unterschiedlichen Kollaborationstypen passen wir bestehende Modelle zur Teamperformance auch bezogen auf die Prozess-Faktoren an. Die Art der Kollaboration (bestehend aus physikalischer Distanz, zeitlicher Distanz, Bedarf an Kommunikationstechnologie) werden hier als Mediatorvariable konzeptualisiert und als Einflussfaktor auf den Output ergänzt.

Organisation von Gruppendiskussionen – Videostudien im problemorientierten Wirtschaftsunterricht

Christin Siegfried
Georg-August-Universität Göttingen

Theoretischer Hintergrund

Ökonomische Kompetenz ist als anerkannter Teil der Allgemeinbildung in den letzten Jahren im Schulcurriculum verstärkt verankert worden und soll Lernenden u.a. dazu befähigt „Ansätze zur Lösung nationaler, europäischer und globaler Herausforderungen … (zu) verstehen und sich mit unterschiedlichen Lösungsperspektiven auseinander(zu)setzen“ (Hessisches Kultusministerium, 2016, S. 10).

Der folgende Beitrag untersucht Gruppendiskussionen im Rahmen eines problemorientierten Wirtschaftsunterrichts (siehe hierzu Reusser, 2005). Diesem Unterrichtssetting liegt die Annahme zugrunde, die Anforderungen ökonomischer Kompetenz, wie die Mehrperspektivität und die Auseinandersetzung mit unterschiedlichen Handlungsoptionen durch Gruppendiskussionen geeignet adressieren zu können (Eberle, 2006; Retzmann et al., 2012). Befunde verweisen darauf, dass Gruppendiskussionen, wie sie im Problemorientierten Unterricht verankert sind, neben positiven Effekten für das Lernen (Chi & Menekse, 2019, S. 715), sich auch günstig auf die Lernmotivation sowie das kognitive Engagement von Lernenden auswirken (Seifried, 2005, S. 239). Auch wenn die dargestellten Befunde ein positives Bild kooperativen Arbeitens zeichnet, beschränken sich beispielsweise Gruppendiskussionen im Unterricht nicht nur auf Themen, die den eigentlichen Lerngegenstand betreffen. Zum einen gelingt es nicht immer, sich inhaltlich an der Lernaufgabe auszurichten, sodass auch private Gespräche Teil einer Gruppendiskussion im Unterricht sein können (Nolting, 2017, S. 15, 18). Darüber hinaus entfallen in einer Gruppendiskussion auch Gesprächsbestandteile auf die Aushandlung von Interaktionsregeln, der Planung und Überwachung des Diskussions- bzw. Lernprozesses (Stempfle & Badke-Schaub, 2002, S. 60; Wuttke, 2005, S. 18). Die damit verbundene kontinuierliche Koordination und Organisation der Gruppendiskussion ist notwendig, um die gewonnen Informationen stetig zu überprüfen, zu strukturieren und in eine kognitiv leichter zu verarbeitende Form zu bringen (Stempfle & Badke-Schaub, 2002, S. 60; Schiefele und Wild, 1994, S. 186). Gelingt dies bspw. nicht, kann das leicht Auswirkungen auf die Qualität der gesamten Gruppendiskussion haben, bis hin dazu, dass es nicht gelingt, eine Gruppendiskussion zu führen (Nolting, 2017, S. 18-19).

Fragestellung

1. Welche Gruppenorganisationsprozesse lassen sich für Gruppen in Präsenz, hybrid und online identifizieren?

2. Welche Unterschiede ergeben sich für Gruppenorganisationsprozesse in den unterschiedlichen Settings?

Methode

Ziel des Vortrags ist es Gemeinsamkeiten und Unterschiede dieser Organisationsprozesse in Gruppendiskussion in drei unterschiedlichen Settings nämlich in Präsenz, hybrid und online zu analysieren. Hierfür wurde eine problemorientierte Unterrichtseinheit entwickelt und im Jahr 2017 in Präsenz mit 5 Klassen (90 Lernenden in 16 Gruppen) der hessischen Oberstufe, im Jahr 2019 hybrid mit 4 Klassen (60 Lernenden in 15 Gruppen) und im Jahr 2021 mit 2 Klassen (50 Lernende in 9 Gruppen) durchgeführt. Der gesamte Lehr- und Lernprozess jeder Gruppe (90 Minuten) wurde mit Videos aufgezeichnet und mit Fragebögen (demografische Daten, aktuelle Motivation, Fachinteresse) und Wissenstests vor, direkt nach dem Unterricht und 6 Wochen nach dem Unterricht begleitet. Alle Videos wurden transkribiert und nach der Durchführung einer Kodierschulung kodiert (Interrater-Realibiltät:.72-.83)

Ergebnisse

Die Ergebnisse verweisen zunächst auf einen signifikanten Lernzuwachs über die Unterrichtseinheit, der auch 6 Wochen danach noch anhält ($F(2,162)=5.362$, $p=.007$, $\eta 2=0.15$).

Weiterhin macht die Analyse der vorliegenden Daten deutlich, dass die Organisation der Gruppendiskussionen eine große Rolle spielt. Sie nimmt nicht nur - relativ gesehen - viel Zeit der Gruppendiskussion in Anspruch, sondern hat (insbesondere die Planung und Regulation)– bis zu einem gewissen Grad – einen positiven Effekt auf den Wissenszuwachs der Lernenden. In den Gruppen die hybrid und digital zusammengearbeitet haben zeigen sich erwartungskonform Gruppenorganisationsprozesse, die sich auf technische Schwierigkeiten zurückführen lassen, Darüber hinaus lassen sich allerdings keine signifikanten Unterschiede in den Gruppenorganisationsprozessen zwischen den Gruppen, die in Präsenz, hybrid oder online zusammengearbeitet haben ausmachen.

Effektiveres Zusammenarbeiten, neue Leute kennenlernen“ – Empirische Studien über die Akzeptanz algorithmischer Gruppenformation bei (virtuellen) Lerngruppen

Henrik Bellhäuser[1], Anna Leitz[1], Lea Dreifert[2], René Röpke[3]
[1]Johannes Gutenberg Universität Mainz, [2]Kienbaum Institut, [3]RWTH Aachen University

Theoretischer Hintergrund

Kooperatives Lernen ist eine effektive Lernstrategie und wird in der online-gestützten Hochschullehre vielfach eingesetzt. Wenig Beachtung findet dabei üblicherweise die Gruppenformation, also der Auswahlprozess für die Mitglieder der Lerngruppen, obwohl die Zusammensetzung von Lerngruppen maßgeblich zu ihrem Erfolg beiträgt (Bell, 2007; Halfhill et al., 2005). In der Hochschullehre kommen dabei typischerweise zwei Verfahren zum Einsatz: Selbstselektion durch die Studierenden oder Zufallsselektion durch die Dozierenden. Beide Möglichkeiten haben wesentliche Nachteile, optimieren in keiner Weise die Zusammenstellung für das angestrebte Lernziel und nutzen nicht das Potential digitaler Unterstützung.

Neuere Ansätze ermöglichen kriterienbasierte Gruppenformation mittels Algorithmus (Bellhäuser et al., 2018; Müller et al., 2022), was insbesondere für virtuelle Lerngruppen vielversprechend scheint. Allerdings setzt dies die Bereitschaft der Lernenden voraus, die Entscheidung über die Gruppenformation an den Algorithmus abzugeben. Eine qualitative Studie (Schumacher & Kuzilek, 2022) konnte bereits tendenziell positive Einstellungen feststellen. Gemäß dem UTAUT-Modell (Unified Theory of Acceptance and Use of Technology; Venkatesh et al., 2003) ist die Nutzung neuer Technologien abhängig von den Erwartungen bezüglich Nutzen und Aufwand der Technologie, sozialen Einflüssen und weiteren erleichternden Faktoren.

Fragestellung

Basierend auf diesem Modell untersuchen wir in zwei empirischen Studien die Einstellung von Studierenden gegenüber algorithmischer Gruppenformation: Wie groß ist die Akzeptanz gegenüber algorithmischer Gruppenformation? Welche Faktoren (wie z.B. Persönlichkeitseigenschaften oder Soziodemographie) beeinflussen diese Akzeptanz? Wie entscheiden sich Studierende, wenn Sie die Wahl zwischen algorithmischer Gruppenformation und selbstgebildeten Gruppen haben? Und: Welche Faktoren beeinflussen diese Wahl?

Methode und Ergebnisse

Studie 1 wurde durchgeführt mit N=478 Studierenden aus mehreren deutschen Universitäten, Hochschulen für angewandte Wissenschaften und einer privaten Universität. In einem querschnittlichen Design füllten die Versuchspersonen Fragebögen aus zu ihrer Persönlichkeit (BFK-I; Rammstedt & John, 2005; fünf Skalen, insgesamt 21 Items; Ω = .71-.85), zu Teamorientierung (Campion et al., 1993; drei Items; Ω = .89), einen eigens konstruierten Fragebogen zur Akzeptanz algorithmischer Gruppenformation (vier Items; Ω = .86) sowie soziodemographische Angaben.

Die Ergebnisse zeigen, dass insgesamt in der Stichprobe eine positive Einstellung gegenüber algorithmischer Gruppenformation vorherrschte (M = 4.26 (SD = 1.19) auf einer sechsstufigen Likertskala). Erwartungskonform zeigten sich Korrelationen mit Offenheit (r = .13), Extraversion (r = .13) und sozialer Verträglichkeit (r = .12). Geschlecht, Alter, Semesteranzahl, Studienfach und Herkunftsland beeinflussten hingegen nicht die Akzeptanz. In offenen Fragen wurde vor allem effizientere Gruppenarbeit und die Möglichkeit neue Menschen kennenzulernen als potentielle Vorteile einer algorithmischen Gruppenformation genannt.

In Studie 2 wurden N=393 Informatikstudierende in einer einführenden Vorlesung mit Tutorium vor die Wahl gestellt, ob sie selbst eine Lerngruppe bilden (manuelle Gruppenformation) oder an der algorithmischen Gruppenformation teilnehmen möchten. Neben den Messinstrumenten aus Studie 1 wurden zusätzlich Outcome-Maße der Gruppenarbeit erhoben (Gruppenperformance, Zufriedenheit).

Es zeigte sich, dass sich n=132 (33,6%) der Studierenden für die algorithmische Gruppenformation entscheiden. Diese Personen waren signifikant älter (MAlgorithmus=23.2, SD=3.78 vs. MManuell=21.3, SD=2.28), hatten schlechtere Abiturnoten (MAlgorithmus=2.21, SD=.61 vs. MManuell=1.90, SD=.58) und waren weniger gewissenhaft (MAlgorithmus=3.59, SD=.57 vs. MManuell=3.85, SD=.41). Gruppenleistung (MAlgorithmus=90.5, SD=14.8 vs. MManuell=97.4, SD=12.9) sowie Zufriedenheit (MAlgorithmus=4.3, SD=1.4 vs. MManuell=5.1, SD=1.0) fielen ebenfalls zugunsten der manuell gebildeten Gruppen aus.

Aufgrund des quasiexperimentellen Designs können Leistung und Zufriedenheit nicht kausal auf den Gruppenformationsalgorithmus zurückgeführt werden. Die deutlich ungünstigeren Ausgangswerte der algorithmischen Gruppen schon zu Beginn der Studie lassen bereits schlechtere Leistung und niedrigere Zufriedenheit erwarten – es wäre sogar vorstellbar, dass diese Personengruppe mit selbstgewählten Lerngruppen noch schlechter abgeschnitten hätte.

Insgesamt zeigte sich in beiden Studien, dass Studierende eine durchaus positive Einstellung gegenüber algorithmischer Gruppenformation haben und diese auch tatsächlich in realen Auswahlsituationen wählen. Allerdings scheinen mehrere negative Leistungsprädiktoren mit dieser Wahl assoziiert zu sein, weshalb diese Personengruppe einen besonderen Unterstützungsbedarf zu haben scheint. Algorithmische Gruppenformation kann folglich eine sinnvolle Ergänzung, gerade in virtuellen Gruppenarbeitsszenarien darstellen.

E-17: Eye Tracking als Methode zur Erfassung professioneller Unterrichtswahrnehmung

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* T03 R04 D10

Eye Tracking als Methode zur Erfassung professioneller Unterrichtswahrnehmung

Chair(s): **Andreas Gegenfurtner** (Universität Augsburg, Germany), **Kathleen Stürmer** (Universität Tübingen)
*Diskutant*in(nen):* **Rebekka Stahnke** (IPN Kiel)

Eine wachsende Zahl an Studien in der empirischen Bildungsforschung nutzt Eye Tracking, um die visuellen Wahrnehmungsprozesse von Lehrpersonen im Klassenzimmer zu untersuchen und daraus Rückschlüsse auf die professionelle Unterrichtswahrnehmung von Lehramtsstudierenden und erfahrenen Lehrkräften abzuleiten (für Reviews dieser Arbeiten, siehe Grub et al., 2020; König et al., 2022). Mit Eye Tracking lassen sich etwa die Blickbewegungen einer Lehrperson erfassen, die Zahl und Dauer der Fixationen auf bestimmte Schüler*innen, die Aufmerksamkeitsverteilung im Klassenzimmer oder auch die Dauer bis zum ersten Blick auf für das Unterrichtsgeschehen kritische Ereignisse (Cortina et al., 2015; Grub et al., 2022; Huang et al., in press; Kosel et al., 2021; Stahnke & Blömeke, 2021; Stürmer et al., 2017). Auf theoretischer Ebene ist die Verarbeitung visueller Aufmerksamkeits- und Wahrnehmungsprozesse häufig als *noticing* oder *attending* konzeptualisiert—in Tandem zu wissensbasierten Erklärprozessen dessen, was visuell wahrgenommen wird (siehe etwa theoretische Modelle von Blömeke et al., 2015; Gegenfurtner et al., 2022; Seidel & Stürmer, 2014; Seidel et al., in Druck; van Es & Sherin, 2021). Zur empirischen Testung dieser theoretischen Annahmen zu professioneller Unterrichtswahrnehmung und visueller Expertise leistet Eye Tracking einen wichtigen Beitrag.

Gleichzeitig sind Arbeiten mit Eye Tracking nicht ohne Herausforderung. Orquin und Holmqvist (2018) weisen auf eine Reihe von Risiken hin, die die Reliabilität und Validität von Eye Tracking-Studien beeinträchtigen können, etwa die Vergleichbarkeit, Datenqualität, unterschiedliche Metriken und Parametereinstellungen und die Interpretation der Daten, wenn Eye Tracking als alleiniges Erhebungsverfahren genutzt wird. In diesem Symposium werden einige dieser Risiken thematisiert. Beitrag 1 beispielsweise adressiert die Vergleichbarkeit von Ergebnissen im mobilen Eye Tracking, die durch die Nutzung simulierter und standardisierter Unterrichtssituationen erhöht werden kann und dadurch auch eine genaue Analyse der Aufmerksamkeitsverteilung in verschiedenen Unterrichtsphasen ermöglicht. Beitrag 2 nutzt einen Multi-Method-Ansatz, um die Einflüsse individueller Personenmerkmale wie Einstellungen und Überzeugungen auf das Blickverhalten zu untersuchen; dadurch werden Erkenntnisse gewonnen, die mit Eye Tracking alleine nicht möglich wären. Ähnlich arbeitet Beitrag 3 im Rahmen einer Mixed-Method-Studie, um die Augenbewegungsdaten von Noviz*innen und Expert*innen im Klassenzimmer mit qualitativen Verbaldaten zu triangulieren; gerade diese Datentriangulation erlaubt vertiefte Einblicke in die Gründe und Motive hinter visuellen Aufmerksamkeitsprozessen. Beitrag 4 schließlich berichtet über Ergebnisse einer Studie mit Lehramtsstudierenden, bei der die Test-Retest-Reliabilität und einige wichtige Parameter in der Analyse von Eye-Tracking-Daten gezielt manipuliert wurden. Das Symposium endet mit der Diskussion einer renommierten Unterrichtsforscherin über die Implikationen dieser vier Beiträge zur Integration von Eye Tracking in das Methodenrepertoire der empirischen Unterrichtsforschung.

Beiträge des Symposiums

Alle im Blick? – Zum Einsatz kognitiver Prompts beim Noticing angehender Lehrkräfte in einer standardisierten Unterrichtssimulation

Leonie Telgmann, Katharina Müller
Leibniz Universität Hannover

Theoretischer Hintergrund

Die Klassenführung gilt als bedeutende Dimension von Unterrichtsqualität (Praetorius et al., 2018). Als zentrale Voraussetzung für die Umsetzung von Klassenführungsstrategien, müssen Lehrkräfte auf Grundlage kognitiver Skripts in der Lage sein, ihre visuelle Aufmerksamkeit auf relevante Situation zu richten und diese entsprechend zu interpretieren (Wolff et al., 2021). Erstere Fertigkeit, die wissensbasierte Aufmerksamkeitssteuerung, wird im Konzept der professionellen Unterrichtswahrnehmung als Noticing bezeichnet (Seidel, 2022). Zur Erfassung dieser liefern Blickbewegungen sowohl während der Videobeobachtung (on-action: z.B. Stahnke & Blömeke, 2021) als auch während des Unterrichtens (in-action: z.B. Stürmer et al., 2017) objektive Maßzahlen. (Mobile) Eye-Tracking Studien zeigen, dass Noviz:innen im Vergleich zu Expert:innen ihre Aufmerksamkeit weniger auf die Schüler:innen richten (z.B. Wolff et al., 2016, Huang et al., 2021) und Schwierigkeiten haben, ihre Fixationen gleichmäßig auf alle Schüler:innen im Klassenzimmer zu verteilen (z.B. Cortina et al. 2015, van Bogert et al., 2014). Um Noviz:innen angemessen bei der Entwicklung ihrer Klassenführungsfähigkeiten zu unterstützen, finden sich experimentelle Eye-Tracking Studien, die Effekte von Instruktionen (z.B. Prompts) auf das Noticing bei der Videobeobachtung (on-action) untersuchen (Grub et al., 2022). Unklar bleibt, ob und welche Maßnahmen helfen, um dezidiert wissensbasierte Aufmerksamkeitsprozesse während des Unterrichtens (in-action) besser zu steuern. Ziel dieser quasi-experimentellen Studie ist es zu untersuchen, welchen Effekt kognitive Prompts auf die wissensbasierte Aufmerksamkeitssteuerung während des Unterrichtens (in-action) in verschiedenen Lehr-Lehrsettings haben.

Forschungsfrage

Wie unterscheiden sich zwei Experimentalgruppen (Prompt vs. kein Prompt) in ihrer Aufmerksamkeitsverteilung im Rahmen einer standardisierten Unterrichtssimulation in einer (a) lehrer:innenzentrierten Phase, (b) schüler:innenzentrierten Phase und (c) Übergangsphase?

Design und Methodik

Im Sommer 2021 nahmen hierzu N = 24 Lehramtsstudierende im Master (Mage = 26.42, SDage = 4.58, 67 % weiblich) nach einem Seminar zum Thema Unterrichtseffektivität/Klassenführung an einer 20-minütigen standardisierten Unterrichtssimulation teil, wobei eine Gruppe (N=10) vor und während der Simulation zusätzlich ein kognitives Prompting zu evidenzbasierten Klassenführungsstrategien erhielt. Die Simulation gliederte sich in drei vorgegebene Lehr-Lernphasen (a-c), deren Zeitspanne im Nachgang von zwei trainierten Beobachterinnen kodiert wurde (κ >.8). Um Effekte auf wissensbasierte Aufmerksamkeitsprozesse prüfen zu können, wurden reizinduzierte Aufmerksamkeitsprozesse für alle Lehramtsstudierenden auf einem ähnlichen Niveau etabliert, indem sieben Schauspieler:innen als Schüler:innen aktives und passives On- oder Off-Task-Verhalten zu einem vorher festgelegten Zeitpunkt zeigten. Blickbewegungen wurden mit Tobii Glasses Pro 3, 50

Hz aufgenommen. Für die Analyse wurden alle Schüler:innen als Areas of interest (AOIs) definiert und die Blickbewegungen manuell kodiert. Als objektive Maßzahl für die Aufmerksamkeitsverteilung auf die einzelnen Schüler:innen verglichen wir mittels T-Tests den GINI-Koeffizienten (z.B. Cortina et al., 2015, Smidekova et al., 2020) für die Anzahl von Aufenthalten (visit count) und durchschnittliche Dauer (visit duration) in den AOIs von beiden Gruppen während der verschiedenen Lehr-Lernsettings.

Ergebnisse und Diskussion

Die berechneten GINI-Koeffizienten zeigen, dass sich die Aufmerksamkeitsverteilung beider Gruppen (Prompt vs. kein Prompt) in der standardisierten Unterrichtssimulation unterscheiden. Dies gilt sowohl für die Anzahl und Dauer der Aufenthalte in allen drei Lehr-Lehrsettings (a-c). Der größte Effekt zeigt sich in der schüler:innenzentrierten Phase (d= .58). In der Tendenz weist die Prompt-Gruppe hier im Vergleich zur Kein-Prompt-Gruppe durchschnittlich niedrigere GINI Koeffizienten (Anzahl der Aufenthalte) vor (GKPrompt= 0.20; GKKeinPrompt=0.25; pone-sided=.089). Der Prompt-Gruppe scheint es insbesondere in diesem Setting besser zu gelingen, ihre Aufmerksamkeit gleichmäßiger zu steuern, wobei T-Tests keine signifikanten Unterschiede zeigen. Weitere Unterschiede werden im Vortrag dargelegt und diskutiert.

Die Ergebnisse der vorliegenden Studie bieten mit einer standardisierten Unterrichtssimulation neuartige Einblicke in das Blickverhalten von angehenden Lehrpersonen im Rahmen unterschiedlicher Lehr-Lernsettings. Möglich wird die Identifikation von Ansatzpunkten für Unterstützungsmaßnahmen in der Lehrkräftebildung, bspw. durch die hier eingesetzte komplexitätsreduzierte Unterrichtsimulation. Insbesondere das quasi-experimentelle Setting stellt einen ersten Schritt dar, um zu verstehen, ob ein kurzfristiges Training wissensbasierter Aufmerksamkeitsprozesse von angehenden Lehrkräften mithilfe von kognitiven Prompts zielführend erscheint.

Der Blick auf die Lernenden? – Aufmerksamkeitsverteilung von Lehramtsstudierenden in heterogenen Klassenzimmern

Özün Keskin, Andreas Gegenfurtner
Universität Augsburg

Theoretischer Hintergrund

Seidel und Stürmer (2014) definieren professionelle Unterrichtswahrnehmung als die Fähigkeit, relevante Unterrichtssituationen zu erkennen und zu interpretieren. Studien im Bereich der professionellen Unterrichtswahrnehmung werden häufig mit der Eye-Tracking-Technologie durchgeführt, um die Augenbewegungen von Lehrkräften (z. B. die Anzahl und Dauer ihrer Fixationen) während des Unterrichtsgeschehens genau zu beobachten und für weitere Analysen zugänglich zu machen (Goldberg et al., 2021).

Heterogenität ist ein weit gefasster Begriff, der in unterschiedlichen Kontexten und Settings zu finden ist. Im schulischen Kontext ist vor allem Migrationshintergrund eine große Herausforderung. Bislang ist noch unklar, wie Lehrkräfte ihre Aufmerksamkeit auf Schüler*innen mit Migrationshintergrund lenken und diese visuell wahrnehmen. Zweifelsohne impliziert Lehrkraftprofessionalität, dass Lehrkräfte Schüler*innen ohne Voreingenommenheit oder Vorurteile wahrnehmen, um eine hohe Unterrichtsqualität für alle Schüler*innen unabhängig von ihrer Herkunft zu gewährleisten (Hofer,

2009). Studien scheinen jedoch darauf hinzudeuten, dass Lehrkräfte eher dazu neigen, Schüler*innen mit Migrationshintergrund weniger zu respektieren als Schüler*innen ohne Migrationshintergrund (Vieluf & Sauerwein, 2018).

Das Equity-Modell von van Es und Kolleg*innen (2022) betont, dass Lehrkräfte Schüler*innen als vollständige Person wahrnehmen sollten, ohne sie auf ihren Migrationshintergrund zu reduzieren. Bisherige Studien, die die Einstellung von angehenden Lehrkräften gegenüber Schüler*innen mit Migrationshintergrund untersuchten, zeigen jedoch, dass es noch immer Diskriminierungen zu geben scheint (Kleen et al., 2019).

Forschungsfragen

Aufgrund fehlender Befunde zu Heterogenitätsaspekten im Zusammenhang mit der professionellen Unterrichtswahrnehmung von Lehrkräften in Eye-Tracking-Studien (Keskin et al., 2022) ist beabsichtigt, die folgenden zwei Forschungsfragen zu untersuchen: (a) Wie richtet sich die Aufmerksamkeit von Lehramtsstudierenden auf Schüler*innen mit und ohne Migrationshintergrund? (b) Welche Einstellung haben Lehramtsstudierende gegenüber Schüler*innen mit Migrationshintergrund im Vergleich zu Schüler*innen ohne Migrationshintergrund? Bei der Beantwortung der Fragen wird davon ausgegangen, dass angehende Lehrkräfte ihre Aufmerksamkeit ungleichmäßig verteilen und der Blick sich mehr auf Schüler*innen ohne Migrationshintergrund richtet (Vieluf & Sauerwein, 2018).

Methode

83 Lehramtsstudierende der Universität Augsburg (Alter M=21,37; SD= 2,90; 79,5% weiblich) nahmen an der Studie teil. Die Teilnehmenden betrachteten Unterrichtsvideos, das Interaktionen zwischen einer Lehrkraft und Schüler*innen zeigte. Während der Betrachtung der Videos nahm ein Tobii Pro Spectrum Eye-Tracker die Augenbewegungen der Teilnehmenden auf. Anschließend wurden die Teilnehmenden qualitativ interviewt und gebeten, einen standardisierten Fragebogen auszufüllen. (a) Das qualitative Interview enthielt Fragen zur Fähigkeit der Lehrkraft im Video, mit den Schüler*innen mit Migrationshintergrund umzugehen und zur selbsteingeschätzten Fähigkeit der Studienteilnehmenden, Schüler*innen mit Migrationshintergrund in einer allgemeinen Unterrichtssituation (z.B. während des Praktikums) wahrzunehmen. (b) Der Fragebogen enthielt Skalen zur Messung der expliziten Einstellung gegenüber Schüler*innen mit Migrationshintergrund (Gefühlsthermometer nach Alwin, 2007), zur multikulturellen Überzeugung (Hachfeld et al., 2012) und zur sozialen Erwünschtheit (Satow, 2012). Vor der Datenerhebung wurde eine statistische Power-Analyse mit dem Softwareprogramm G*Power v3.1 durchgeführt. Für eine lineare multiple Regression mit einer Alpha-Fehlerwahrscheinlichkeit von 0.05 und einer Aussagekraft von 0.90 betrug die empfohlene Gesamtstichprobengröße 77. Unsere Stichprobe von 83 Teilnehmenden übertraf diese Empfehlung, was darauf hindeutet, dass die Stichprobengröße ausreichend groß war und eine angemessene statistische Aussagekraft aufweist.

Ergebnisse und Diskussion

Auf Basis einer explorativen Analyse der qualitativen Daten lässt sich herausstellen, dass eine unterschiedliche Wahrnehmung von Heterogenität im Klassenzimmer vorliegt. Für die Analyse und

Kodierung der qualitativen Daten wird die Software MaxQDA 2022 Analytics Pro verwendet. Nach der Aufbereitung der Blickbewegungsdaten wird sich die Auswertung der Ergebnisse auf die Anzahl und Dauer der Fixationen auf Schüler*innen mit und ohne Migrationshintergrund konzentrieren. Zusätzlich werden die Blickbewegungsdaten mit den Einstellungsmessungen korreliert und mit der qualitativen Auswertung der Interviewdaten trianguliert. Die Ergebnisse werden auf der GEBF-Jahrestagung vorgestellt und können erste Einblicke in das Blickverhalten und die professionelle Wahrnehmung von Lehramtsstudierenden im heterogenen Klassenzimmer liefern.

Wahrnehmung von Unterrichtsstörungen bei (angehenden) Lehrkräften: Ein mixed methods-Ansatz

Ann-Sophie Grub[1], Roland Brünken[1], Doris Lewalter[2], Antje Biermann[1]
[1]Universität des Saarlandes, [2]Technische Universität München

Die professionelle Unterrichtswahrnehmung von Lehrkräften gilt als eine bedeutsame Voraussetzung für unterrichtliches Handeln (Blömeke et al., 2015). Um Situationen schnell wahrzunehmen und diese als relevant für Unterricht zu interpretieren, ist ein effektives Monitoring-Blickverhalten notwendig, welches sich durch viele und kurze Fixationen auszeichnet (Wolff, 2016; Huang et al., 2021). In bisherigen Studien konnte gezeigt werden, dass in der Regel Expert*innen über ein solches Blickverhalten verfügen (vgl. Grub et al., 2020), einige Studien können diese Expertise-Unterschiede jedoch nur teilweise replizieren (Grub et al., 2022, Shinoda et al., 2021, Stahnke & Blömeke, 2021, Wyss et al., 2021). In Hinblick auf die kritische Diskussion der Eye-Mind-Hypothese (Just & Carpenter, 1980) wird eine Triangulation, also die gemeinsame Analyse von Eyetracking-Daten und Verbaldaten empfohlen (Orquin & Holmqvist, 2018), um die dahinterliegenden kognitiven Prozesse abbilden zu können. Obwohl in einigen Studien beide Datenquellen genutzt wurden (Muhonen et al., 2021; 2022; Pouta et al., 2021, Wolff et al., 2015), wurde eine direkte Triangulation kaum durchgeführt. Eine Ausnahme stellt unseres Wissens die Studie von Wyss et al. (2021) dar, in der bei einem Video Eyetracking-Daten und die retrospektive Kommentierung des Gesehenen direkt verknüpft wurden. Wyss und Kollegen konnten zwar Expertise-Unterschiede im Monitoring-Blickverhalten zeigen, allerdings gingen diese Unterschiede nur auf diejenigen Expert*innen zurück, die das kritische Ereignis (Critical Incident, CI) im Video in den Laut-Denken-Protokollen genannt hatten. Anzumerken ist, dass in der Studie lediglich ein Video mit einem kritischen Ereignis verwendet wurde, außerdem nahmen nur sechs Teilnehmer*innen das Ereignis überhaupt war.

Die vorliegende Studie bringt Blickbewegungs- und retrospektive Laut-Denken-Daten von zwei Videos mit jeweils mehreren CI zusammen. Überprüft werden soll erstens die Frage, ob die Ereignisse, die während des Betrachtens eines Klassenraumvideos als relevant eingeschätzt werden (Tastendruck), auch im stimulated recall genannt werden. Zweitens sollen Unterschiede im Blickverhalten (Anzahl und Dauer der Fixationen, Zeitpunkt bis zur ersten Fixation des CI) überprüft werden zwischen (a) Expert*innen und Noviz*innen sowie (b) zwischen Personen, die die CI in den Videos wahrgenommen haben (Responder) und denen, die dies nicht taten (Non-Responder).

An der Studie nahmen N=26 berufstätige Lehrkräfte (Berufserfahrung > 10 Jahre) und 26 Lehramtsstudierende teil. Die Teilnehmer*innen bekamen insgesamt sechs kurze Videos gezeigt, von denen im Beitrag zwei genauer analysiert werden. In diesen Videos waren zwei bzw. drei

Unterrichtsstörungen zu sehen. Die Teilnehmer*innen sollten beim Betrachten der Videos eine Taste drücken, wenn etwas Relevantes wahrgenommen wurde. Die Blickbewegungsdaten wurden aufgezeichnet. Im Anschluss wurden die Teilnehmer*innen gebeten, zu kommentieren, was sie im Video gesehen haben (stimulated recall).

In Hinblick auf die erste Fragestellung zeigte sich, dass die CI signifikant häufiger in den Laut-Denken-Protokollen genannt wurden, als mit einem Tastendruck auf sie reagiert wurde. Hinsichtlich Frage zwei fanden sich keine Expertise-Unterschiede. Allerdings zeigten die Responder ein professionelleres Monitoring-Blickverhalten als die Non-Responder: Sie hatten signifikant häufigere und kürzere Fixationen auf den relevanten AOIs, außerdem haben sie diese auch signifikant schneller wahrgenommen. Die Ergebnisse zeigen sich sowohl auf CI- als auch auf Videoebene.

Die Ergebnisse verdeutlichen zum einen die Relevanz der Methoden-Triangulation: Mit Hilfe von Laut-Denken-Protokollen können die Blickbewegungsdaten um wichtige Informationen ergänzt und ggf. korrigiert werden. Darüber hinaus deuten die Ergebnisse darauf hin, dass nicht die Expertise i.S. von Berufserfahrung allein professionelles Blickverhalten erklären kann, sondern dass auch Lehramtsstudierende in der Lage sind, effektives Monitoring bei der Betrachtung von Unterrichtsvideos anzuwenden. In zukünftigen Studien sollte untersucht werden, welche Determinanten hierfür eine Rolle spielen (z.B. Professionswissen, Selbstwirksamkeit, Überzeugungen). Eine Limitation der Studie besteht in der Nutzung kurzer Videos, was die ökologische Validität der Ergebnisse stark einschränkt. Ein Fokus zukünftiger Forschung kann daher in der Nutzung virtueller Klassenzimmer in Verbindung mit mobilem Eyetracking liegen, womit bei hoher interner Validität auch die ökologische Validität verbessert werden kann.

Wissenschaftliche Gütekriterien in Eye-Tracking Studien mit Lehrpersonen: Herausforderungen und Lösungsansätze

Christian Kosel, Ricardo Böheim, Tina Seidel
Technische Universität München

Einleitung

Die Anzahl an Eye-Tracking Studien in der Lehrer:innenforschung haben in den letzten Jahrzehnten enorm zugenommen (Gegenfurtner, 2019), da die notwendigen Geräte immer günstiger in der Anschaffung und einfacher zu bedienen geworden sind (Orquin & Holmqvist, 2018). Beispielsweise nutzen Wissenschaftler:innen vermehrt Eye-Tracking dazu, die professionelle Wahrnehmung von Lehrkräften im Unterricht zu erfassen (Halszka et al., 2017). Allgemein gesprochen, ist die Zunahme der Eye-Tracking Studien jedoch nicht mit einem wachsenden Bewusstsein für die Methodik einhergegangen. In diesem konzeptionellen und mit empirischen Daten unterfütterten Beitrag skizzieren wir fünf zentrale Herausforderungen bei Eye-Tracking Studien, deren Einfluss auf die in der statistischen Testtheorie entwickelten Hauptgütekriterien (hier Reliabilität und interne/externe Validität) (Przyborski & Wohlrab-Sahr, 2014) bisher nur unzureichend untersucht wurden.

Die Reliabilität in Eye-Tracking Studien ist gefährdet, da eine Vielzahl der verfügbaren Studien ihre Ergebnisse auf einen critical trial beschränken. Es ist jedoch anzunehmen, dass auch in Eye-Tracking Studien Messungen mit einem Messfehler behaftet sind (quantifizierbar mit der Test-Retest-Reliabilität; Farris-Trimble & McMurray, 2013). Im Hinblick auf die interne Validität fokussieren wir auf

den Einfluss der erhobenen Datenqualität (Trackingratio), Größe und Form der Areas of Interest (AOIs) und variierende Geschwindigkeitsschwellen (in °/s) des Algorithmus zur Identifizierung von Fixationen und Sakkaden. Die externe Validität ist in Eye-Tracking Studien insbesondere durch geringe Stichprobengrößen gefährdet. Viele der relevanten Studien im Feld haben einen Stichprobenumfang von <100 Probanden, die in der Regel nur einen Versuchsdurchlauf absolvieren. Die Generalisierung der erhobenen Augenbewegungen für eine Gesamtpopulation (z.B. Lehrkräfte) oder andere Stimuli sind daher oftmals eingeschränkt (Orquin & Holmqvist, 2018).

Fragestellungen

Reliabilität

1. Wie hoch fällt die Retest-Reliabilität bei zwei Versuchsdurchgängen aus?

Interne Validität

Inwiefern verändern sich die Ergebnisse,

2a. …wenn die zugrundeliegende Stichprobe hinsichtlich der Datenqualität bereinigt wird?

Variierender Parameter: Datenbereinigung (Trackingratio ≤ .90, reduziertes n = 53) vs. volle Stichprobe N = 68.

2b. …wenn AOIs in Größe und Form variieren?

Variierende Parameter: (a) Konturierte AOIs vs. (b) grobflächige, rechteckige AOIs vs. (c) auf das Gesicht begrenzte AOIs.

2c. …wenn die Geschwindigkeitsschwellen (in °/s) zur Identifizierung von Fixationen und Sakkaden verändert werden?

Variierende Parameter: 30°/s bis 150°/s, 10°/s Schritte.

Externe Validität

3. …wenn aus der Gesamtstichprobe (N=68) per Zufallsauswahl Stichproben mit alternierender Stichprobengröße gezogen wird (n=5, n=10, n=15 usw.).

Methode

Datengrundlage dieses Beitrags bildet ein Eye-Tracking Experiment mit insgesamt 68 Lehramtsstudenten:innen, in dem ein 11-minütiger Videoausschnitt einer typischen Unterrichtssequenz in Mathematik (8. Klasse, Gymnasium) gezeigt wurde (Seidel et al., 2020). Im Videoausschnitt sind insgesamt 22 Schüler:innen durchgehend sichtbar. Die Blickbewegungen wurden mit einem stationären SMI RED 500hz Eye-Tracker erhoben. Die zur Analyse notwendigen AOIs wurden manuell um fünf Schüler:innen gezeichnet. Als Metrik wurde die Anzahl der Fixationen gewählt. Die für diesen Beitrag notwendige statistische Modellierung umfasst Korrelationsanalysen (1), Varianzanalysen (2a-c) und Äquivalenztests (3).

Zentrale Ergebnisse

(1) Das Ergebnis der Korrelationsanalyse zeigt eine mittlere Retest-Reliabilität mit einer Korrelation von r =.81.

(2)

a) Nicht signifkant.

b) Auf die rechteckigen AOIs entfielen signifikant mehr Fixationen als auf den beiden anderen Varianten, Greenhouse-Geisser-F(1.40, 84.40) = 383.19,p < .001, $\eta p2$= .86. Der Unterschied zwischen den rechteckigen AOIs (422 Fixationen) und den konturierten AOIs (296 Fixationen) beträgt 126 Fixationen (Δ von 29,85%).

c) Variierende Geschwindigkeitsschwellen (°/s) haben einen großen Einfluss auf die Anzahl der Fixationen und damit auf die Interpretation der Ergebnisse. Bei einer Schwelle von 30°/s, werden 18% mehr Fixationen klassifiziert als bei einer Konfiguration mit einer Schwelle von 50°/s.

(3) Ab n=55 waren die Ergebnisse von allen zufällig gezogenen Stichproben mit denen der Gesamtstichprobe (N = 68) äquivalent. Mit n=20 sind nur noch 64,3% der Ergebnisse äquivalent.

Die Ergebnisse verdeutlichen, dass eine transparente und einheitliche Erhebung und Auswertung von Eye-Tracking Daten zwingend erforderlich ist. Zur Diskussion stehen Lösungsansätze wie Panelerhebungen und Implemtierung standardisierter Flow-Charts.

E-18: Lehrer-Schüler-Beziehungen: Von der Generalisierbarkeit positiver Befunde

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* R11 T00 D01

Lehrer-Schüler-Beziehungen: Von der Generalisierbarkeit positiver Befunde

Chair(s): **Valentin Emslander** (University of Luxembourg, Luxembourg), **Doris Holzberger** (Centre for International Student Assessment (ZIB), TUM School of Social Sciences and Technology, Technical University of Munich, Germany)
*Diskutant*in(nen):* **Benjamin Fauth** (Institut für Bildungsanalysen Baden-Württemberg (IBBW), Stuttgart, Deutschland; Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Germany)

Zahlreiche Studien zeigen Zusammenhänge zwischen positiven Lehrer-Schüler-Beziehungen und Ergebnisvariablen im schulischen Kontext (z.B. Noten, wahrgenommene Leistung, soziales Verhalten, Lernmotivation; Givens Rolland, 2012; Korpershoek et al., 2016). Hierbei werden Lehrer-Schüler-Beziehungen ganz unterschiedlich konzeptualisiert: Während die Lehrer-Schüler-*Interaktion* vor allem den praktischen Austausch zwischen Lehrkraft und Schüler*in betrachtet (z.B. Beteiligung am Unterrichtsgespräch), verbindet die konstruktive Unterstützung durch die Lehrkraft unterrichtliche und inhaltliche Unterstützung mit der Wärme und Qualität der Lehrer-Schüler-*Beziehung* (vgl. Decristan et al., 2022; Endedijk et al., 2021; Hamre & Pianta, 2001). Beide Aspekte sind bedeutend für den schulischen Lernerfolg (Endedijk et al., 2021).

Vor dem Hintergrund zahlreicher Befunde stellt sich die Frage der Generalisierbarkeit von Effekten positiver Lehrer-Schüler-Beziehungen und -Interaktionen: Profitieren alle Schüler*innen gleichermaßen von emotionaler Nähe und Freundlichkeit der Lehrkraft? Oder zeigen sich Unterschiede, beispielsweise zwischen Jungen und Mädchen oder Schüler*innen verschiedenen Alters, inwieweit sich positive Lehrer-Schüler-Beziehungen auf Ergebnisvariablen auswirken? Auch die meta-analytische Literatur ist sich in diesen Fragen teilweise uneins (vgl. Lei et al., 2016; Roorda et al., 2017). Offen bleibt zudem, welche Schüler*innen besondere Nachteile haben bei einer negativen Lehrer-Schüler-Beziehung (Nurmi, 2012). Unser Symposium zeichnet ein umfassendes Bild dieser wichtigen Beziehung und Interaktion im Herzen der schulischen Bildung anhand der Generalisierbarkeit gefundener Effekte:

Der erste Beitrag vergleicht die Rolle der elterliche Unterstützung mit der Unterstützung durch die Lehrkraft auf die akademische Entwicklung der Schüler*innen. Die umfangreiche längsschnittliche Studie verwendet Daten von Haupt-, Real- und Mittelschulen über einen Zeitraum von vier Jahren in der Adoleszenz. Ein Augenmerk liegt auf Unterschieden in den Zusammenhängen zwischen der fünften und achten Klasse.

Der zweite Beitrag untersucht den Einfluss der Fehlerkultur bei der Beteiligung am Unterrichtsgespräch. Mithilfe von Videodaten werden Unterrichtsstunden in Mathematik und Deutsch untersucht. Die Studie untersucht differenzielle Befunde der Zusammenhänge in Abhängigkeit des Selbstkonzepts und Migrationshintergrunds von Schüler*innen.

Ein dritter Beitrag beschäftigt sich mit dem moderierenden Einfluss der Schülerwahrnehmung von konstruktiven Lehrkraftunterstützung in der Sekundarstufe und ihrem Zusammenhang zu mathematischen Selbstkonzept und Interesse. Das besondere Augenmerk dieser Studie liegt auf den Unterschieden in der Motivation von jungen Frauen und Männern für das Fach Mathematik.

Abschließend bietet ein vierter Beitrag einen Überblick über Korrelate von Lehrer-Schüler-Beziehungen in der Literatur. Das systematische Review von Meta-Analysen illustriert Zusammenhänge, Moderatoren und Studienqualität auf dem Gebiet der Lehrer-Schüler-Beziehungen. Besonders werden Alter und Geschlecht der Schüler*innen ebenso wie wer die Lehrer-Schüler-Beziehung einschätzt kontrovers diskutiert.

Beiträge des Symposiums

Die Rolle der Unterstützung von Eltern und Lehrkräften für die akademische Entwicklung in der Adoleszenz: Ergebnisse einer 4-Jahres Längsschnittstudie

Karen Aldrup[1], Lisa Bardach[2], Richard Göllner[2], Benjamin Nagengast[2], Ulrich Trautwein[2]
[1]IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, [2]Universität Tübingen, Hector-Institut für Empirische Bildungsforschung

Theoretischer Hintergrund

Entsprechend theoretischer Modelle, wie der Erwartungs-Wert-Theorie (Eccles et al., 1983) und der Selbstbestimmungstheorie (Ryan & Deci, 2000), stehen leistungsbezogene Motivation, Verhalten und schulischer Erfolg in engem Zusammenhang mit dem sozialen Kontext. Im Einklang damit, konnte die Bedeutung der Unterstützung durch Eltern und Lehrkräfte empirisch bereits vielfach gezeigt werden (Fan & Chen, 2001; Roorda et al., 2017). Insbesondere in der Adoleszenz verändert sich jedoch die Bedeutung unterschiedlicher Interaktionspartner. So stellen Eltern zwar nach dem Übergang in die weiterführende Schule eine stabile Quelle sozialer Unterstützung dar und könnten somit in dieser Phase besonders wichtig sein, durch die zunehmende Hinwendung zu Gleichaltrigen könnten sie jedoch später an Bedeutung verlieren (del Valle et al., 2010; Gniewosz et al., 2012; Grotevant & Cooper, 1986). Mit Blick auf Unterstützung durch Lehrkräfte scheint diese in der weiterführenden Schule in engerem Zusammenhang mit der Leistung und dem Engagement der Schüler:innen zu stehen als im Grundschulkontext (Roorda et al., 2017). Da die Beziehungen zu Lehrkräften in der weiterführenden Schule insgesamt unpersönlicher werden, könnten einzelne positive Beziehungen besonders wirksam sein (Eccles et al., 1993; Hughes & Cao, 2018).

Fragestellung

Vor diesem Hintergrund widmet sich die aktuelle Untersuchung erstens der Frage, inwieweit emotionale und akademische Unterstützung durch Eltern und Lehrkräfte mit zentralen affektiv-motivationalen, behavioralen und kognitiven Indikatoren des schulischen Erfolgs in der Adoleszenz assoziiert sind. Zweitens wird betrachtet, inwieweit sich die Zusammenhänge zwischen der fünften und achten Klasse verändern. Hierbei vermuteten wir, dass der Zusammenhang mit elterlicher Unterstützung abnimmt, wohingegen der Zusammenhang mit der Unterstützung durch Lehrkräfte stabil bleibt.

Methode

Zur Untersuchung dieser Fragen nutzen wir Daten von 3470 Schüler:innen an Haupt-, Real- und Mittelschulen, die an der TRAIN-Studie teilnahmen (Jonkmann et al., 2013). Von der fünften bis zur achten Klasse beantworteten Sie jährlich Fragen zur akademischen und emotionalen Unterstützung durch Eltern und Lehrkräfte sowie zum Selbstkonzept akademischer Fähigkeiten, ihrer Schulidentifikation, Schulzufriedenheit und Absentismus. Zudem bearbeiteten sie standardisierte

Leistungstests in Deutsch und Mathematik. Beim ersten Messzeitpunkt waren die Teilnehmenden im Mittel M = 11.18 Jahre alt (SD = 0.63), 54.6% waren männlich und 33.4% hatten einen Migrationshintergrund.

Die Daten wurden mit längsschnittlichen Strukturgleichungsmodellen (Little, 2013) in Mplus 7.4 ausgewertet. Um die Schachtelung der Schüler:innen innerhalb von Klassen zu berücksichtigen, wurde der TYPE = COMPLEX-Befehl verwendet. Fehlende Werte wurden mithilfe des Full Information Maximum Likelihood-Verfahrens geschätzt.

Ergebnisse und ihre Bedeutung

Ergebnisse latenter State-Analysen zeigten zunächst, dass akademische und emotionale Unterstützung durch Eltern empirisch distinkte Faktoren darstellten, die über den vierjährigen Studienzeitraum eine geringe bis moderate Stabilität aufwiesen. Akademische und emotionale Unterstützung von Lehrkräften formten einen gemeinsamen Faktor, der ebenfalls eine geringe bis moderate Stabilität zeigte. In längsschnittlichen Strukturgleichungsmodellen sagte emotionale Unterstützung durch Eltern eine stärkere Schulidentifikation, eine höhere Schulzufriedenheit, ein positiveres Selbstkonzept akademischer Fähigkeiten, eine bessere Deutsch- und Mathematikleistung und weniger Absentismus vorher. Für die akademische Unterstützung durch Eltern ergab sich – mit Ausnahme von Schulzufriedenheit – ein ähnliches Befundmuster. Unterstützung durch Lehrkräfte war mit einer höheren Schulidentifikation, Schulzufriedenheit und weniger Absentismus verbunden. Die Zusammenhänge blieben über die Zeit stabil und waren insgesamt klein (βs □ .10); insbesondere der Zusammenhang zwischen der elterlichen emotionalen Unterstützung und der Deutsch- und Mathematikleistung (β = .02, p □ .042).

Zusammenfassend zeigen die Befunde, dass Unterstützung durch Eltern und Lehrkräfte unabhängig vom Alter der Schüler:innen bedeutsam für ihre schulische Entwicklung ist und dass sowohl emotionale als auch akademische Aspekte der Unterstützung relevant sind. Folglich sollten beide Interaktionspartner für ihre potenzielle Funktion sensibilisiert werden. Insbesondere in der Aus- und Weiterbildung von Lehrkräften bietet sich ein Ansatzpunkt, um systematisch notwendiges Wissen (z.B. im pädagogisch-psychologischen Bereich; Voss et al., 2015) und Können zu vermitteln (z.B. MyTeachingPartner, Allen et al., 2011).

Welche Rolle spielt die Fehlerkultur bei der Beteiligung am Unterrichtsgespräch? Eine Untersuchung des Haupteffekts und der moderierenden Rolle von Selbstkonzept und Migrationshintergrund

Simon Munk[1], Ricardo Böheim[2], Doris Holzberger[1]
[1]Centre for International Student Assessment (ZIB), TUM School of Social Sciences and Technology, Technical University of Munich, Germany, [2]Technische Universität München, Friedl Schöller-Stiftungslehrstuhl für Pädagogische Psychologie

Theoretischer Hintergrund

Damit gute Lehrer-Schüler-Beziehungen entstehen können, scheinen produktive Lehrer-Schüler-Interaktionen zentral (Scherzinger et al., 2020). In diesem Beitrag werfen wir einen Blick auf eine Facette der Lehrer-Schüler-Interaktion, die Beteiligung am Unterrichtsgespräch. Sie ist bedeutsam für den Lernerfolg der Schüler*innen (Schnitzler et al., 2021). Unter Beteiligung verstehen wir in diesem Beitrag die Schülerverhaltensweisen Melden und Reinrufen sowie das Aufrufen durch die Lehrkraft.

Blickt man auf unterrichtsspezifische Faktoren, so scheint die konstruktive Unterstützung (z. B. Denn et al., 2019) durch die Lehrkraft wichtig für die Beteiligung zu sein. In unserem Beitrag greifen wir nun eine Facette der konstruktiven Unterstützung, die Fehlerkultur, heraus und untersuchen, inwiefern sie eine Rolle für die Schülerbeteiligung am Unterrichtsgespräch spielt. Mit Fehlerkultur können Verhaltensweisen von Schüler*innen und Lehrkräften zusammengefasst werden, die Auskunft über den Umgang mit Fehlern geben (Steuer et al., 2013).

Darüberhinaus analysieren wir, ob die Fehlerkultur für die Beteiligung für alle Schüler*innen gleich bedeutsam ist. Nach dem Aptitude-Treatment-Interaction-Ansatz (Cronbach, 1975) ist anzunehmen, dass für unterschiedliche Schüler*innen auch eine unterschiedlich starke Bedeutung von Fehlerkultur einhergeht. Zwei Schülermerkmale zeigen sich als wichtige Moderatoren für die Bedeutung der konstruktiven Unterstützung der Lehrkraft auf den Lernerfolg: Selbstkonzept (Decristan et al., 2022) und Migrationshintergrund (Seiz et al., 2016). Wir greifen deshalb diese Ergebnisse auf und analysieren, ob die beiden Merkmale auch für den Zusammenhang zwischen Fehlerkultur und Beteiligung einen Unterschied machen. Da für Schüler*innen mit niedrigem Selbstkonzept Fehler-Machen möglicherweise beängstigender ist, könnte für diese Schüler*innen auch die Fehlerkultur eine größere Rolle spielen. Auch Schüler*innen mit Migrationshintergrund, die durchschnittlich eine geringere Leistung (Weis et al., 2018) zeigen und sich weniger zugehörig zu ihren Mitschüler*innen fühlen (Stanat et al., 2019), könnten stärker von einer wertschätzenden Fehlerkultur profitieren.

Fragestellung

Auf Basis dieser theoretischen Annahmen untersuchen wir die folgenden Fragen:

1. Hängen Fehlerkultur, Selbstkonzept und Migrationshintergrund mit Schülerbeteiligung zusammen?

2. Moderiert das Selbstkonzept den Zusammenhang zwischen Fehlerkultur und Schülerbeteiligung?

3. Moderiert der Migrationshintergrund den Zusammenhang zwischen Fehlerkultur und Schülerbeteiligung?

Um die Generalisierbarkeit der Ergebnisse über Unterrichtsfächergrenzen hinweg besser einschätzen zu können, führen wir die Analysen sowohl in Deutsch als auch in Mathematik durch.

Methode

Zur Beantwortung der Forschungsfragen wird die Beteiligung am Unterrichtsgespräch aus Videodaten von jeweils einer Unterrichtsstunde in Deutsch und Mathematik von 20 8. Klassen bayerischer Gymnasien (Jurik et al., 2015) gewonnen. Zusätzlich wurden die N = 387 Schüler*innen (58 % weiblich, Mage = 13.81 Jahre [SD = 0.54]) mit Hilfe eines Fragebogens zu Selbstkonzept, Migrationshintergrund und ihrer Einschätzung der Fehlerkultur befragt. Für die Analysen wurden random-intercept-Modelle mit dem R-Package mitml (Grund, Robitzsch, & Lüdtke, 2021) gerechnet und fehlende Werte mit mdmb (Grund, Lüdtke, & Robitzsch, 2021) geschätzt.

Ergebnisse und ihre Bedeutung

Hinsichtlich Forschungsfrage 1 zeigt sich eine bedeutsame Rolle der Fehlerkultur im Fach Deutsch, nicht jedoch im Fach Mathematik. Das Selbstkonzept hingegen scheint für die Beteiligung am Unterrichtsgespräch nur in Mathematik nicht aber in Deutsch eine signifikante Bedeutung zu haben.

Der Migrationshintergrund zeigt in beiden Fächern keinen bedeutenden Zusammenhang mit der Beteiligung. Hinsichtlich der Forschungsfragen 2 und 3 konnten wir entgegen unserer Annahmen keine signifikanten Unterschiede in der Bedeutung der Fehlerkultur für Schüler*innen mit hohem oder niedrigem Selbstkonzept bzw. mit und ohne Migrationshintergrund feststellen. Dies fügt sich in eine Reihe früherer Ergebnisse, die ebenfalls keine Aptitude-Treatment-Interactions finden konnten (Tetzlaff et al., 2021).

Dieser Beitrag gibt durch die Verknüpfung von Unterrichtsvideographien und Schülereinschätzung zur Fehlerkultur einen bedeutenden Einblick in die Faktoren, die zur Beteiligung am Unterrichtsgespräch führen. Bestehende Befunde werden um das Ergebnis erweitert, dass offenbar fachspezifische stärker als schülerspezifische Gründe einen Unterschied in der Beteiligung machen. Damit zeigen die Analysen einige Stellschrauben auf, die zu produktiven Lehrer-Schüler-Interaktionen und damit auch guten Lehrer-Schüler-Beziehungen beitragen können.

The role of secondary school student perceptions of teacher constructive support for gender differences in motivational outcomes in mathematics

Kaley Lesperance[1], Jasmin Decristan[2], Doris Holzberger[1]
[1]Centre for International Student Assessment (ZIB), TUM School of Social Sciences and Technology, Technical University of Munich, Germany, [2]Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal

Theoretical Background

Male and female students typically differ in their motivational outcomes, such as interest, self-concept and self-efficacy (Wigfield et al., 2002). These differences exist in subjects where, stereotypically, one gender or the other is viewed as more capable. Mathematics is a subject that is stereotypically "male," and where female students usually exhibit lower levels of these motivational factors (Else-Quest et al., 2010; Pajares, 2005). To reduce these differences, it is important to identify what aspects of student classroom experiences could positively promote motivational outcomes. Research shows that instructional quality has an impact on many motivational factors (Burić & Kim, 2020; Scherer & Nilsen, 2016). One aspect that is strongly associated with student motivation is teacher constructive support (Allen et al., 2006; Cornelius-White, 2007). Teacher constructive support refers to instructional and content-related support, as well as the quality of the student-teacher relationship (Decristan et al., 2022; Sliwka et al., 2019). Although research shows that teacher constructive support has a positive effect on students' motivational factors (Fauth et al., 2014), and moderates the relationship between learner characteristics and outcomes (Decristan et al., 2022; Pianta et al., 2008), it is still unclear whether there are differential effects for males and females, i.e., whether teacher constructive support could help reduce gender differences in motivation in secondary school mathematics.

Research Questions

Our following research questions (RQs) are as follows:

1. Do females have significantly lower levels of self-efficacy, self-concept and interest than males in secondary school mathematics?

2. Are student perceptions of teacher constructive support related to students' self-efficacy, self-concept and interest in secondary school mathematics?

3. Do student perceptions of teacher constructive support moderate the relationship between gender and students' self-efficacy, self-concept and interest in secondary school mathematics?

Methods

We used the student questionnaire data from the Teaching and Learning International Survey (TALIS) Germany 2018 Video Study. 1,140 students (mean age = 15) from 50 classes participated. Students completed a pre-test questionnaire, participated in two mathematics teaching units, and then completed a post-test. We used items from the post-test pertaining to the following dependent variables:

• Math self-concept

• Math interest

• Math self-efficacy

We used students' perception of teacher support (instructional/content support) and perception of student-teacher relationship (socio-emotional support) at post-test as indicators of teacher constructive support, and student gender from the pre-test.

We used structural equation models with gender (RQ1), and teacher support and student-teacher relationship (RQ2) as predictors for student motivational factors. For RQ3, we fit two models, one with teacher support and the other model with student-teacher relationship as the moderator for the relationship between gender and motivational outcomes.

Results and Implications

Preliminary analyses showed that girls reported lower mathematics self-concept ($\beta = –.08$, $p < .05$) and lower mathematics self-efficacy ($\beta = –.21$, $p < .001$) than boys. There was no significant effect of gender on mathematics interest. Results also showed that teacher support significantly predicted all student motivational outcomes (self-concept: $\beta = .26$, $p < .001$; interest: $\beta = .37$, $p < .001$; self-efficacy: $\beta = .28$, $p < .001$), as did student-teacher relationship (self-concept: $\beta = .21$, $p < .001$; interest: $\beta = .39$, $p < .001$; self-efficacy: $\beta = .23$, $p < .001$). Preliminary latent moderated SEM did not reveal any moderation effects between gender and teacher constructive support.

This study contributes to the growing evidence that male and female students differ in their motivation factors in mathematics. The non-significant interaction effects indicate that teacher constructive support seems not to compensate for gender differences. However, teacher constructive support is still of great importance for student self-concept, self-efficacy and interest in secondary school mathematics, regardless of gender.

Lehrer-Schüler-Beziehungen und ihre Korrelate: Ein systematisches Review von Meta-Analysen

Valentin Emslander[1], Doris Holzberger[2], Antoine Fischbach[1], Ronny Scherer[3]
[1]Luxembourg Centre for Educational Testing (LUCET) at the University of Luxembourg, Faculty of Humanities, Education and Social Sciences, University of Luxembourg, Luxembourg, [2]Centre for International Student Assessment (ZIB), TUM School of Social Sciences and Technology, Technical University of Munich, Germany, [3]Centre for Educational Measurement at the University of Oslo (CEMO), Faculty of Educational Sciences, University of Oslo, Norway

Theoretischer Hintergrund

Gute Beziehungen zur eigenen Lehrerin können sich positiv auf den Erfolg eines Schülers auswirken. Dieser Effekt kann mit Bowlby's (1982) Bindungstheorie erklärt werden und wird empirisch immer wieder gestützt (z.B. Hamre & Pianta, 2001). Positive Lehrer-Schüler-Beziehungen zeichnen sich durch emotionale Wärme und Nähe aus; negative Aspekte durch Konflikt und Abhängigkeit. So stehen positive Lehrer-Schüler-Beziehungen nicht nur mit akademischen Leistungen in Verbindung, sondern auch mit einer Vielzahl anderer, wünschenswerter Schülerentwicklungen. Zahlreiche Meta-Analysen deuten auf signifikante Zusammenhänge zwischen Lehrer-Schüler-Beziehungen und schulischem Engagement, guten Beziehungen zu Gleichaltrigen, exekutiven Funktionen, allgemeinem Wohlbefinden und der Verringerung aggressiver oder störender Verhaltensweisen hin (Endedijk et al., 2021; Nurmi, 2012; Roorda et al., 2017; Vandenbroucke et al., 2018). Diese Befunde sind jedoch weit verstreut in der Literatur, sodass Forschungslücken unentdeckt bleiben. Auch unterscheiden sich bisherige Überblicksarbeiten in ihren Methoden und den gefundenen Zusammenhängen zwischen Lehrer-Schüler-Beziehungen und Ergebnisvariablen von Schüler*innen. Darüber hinaus ist die Literatur uneindeutig, welche Moderatoren (z.B. Alter oder Geschlecht) diese Beziehungen beeinflussen. Gleichzeitig variiert die Qualität der Meta-Analysen in diesem Feld merklich, was die Interpretation ihrer Ergebnisse erschweren kann.

Fragestellung

Angesichts dieser Forschungslücken haben wir die meta-analytische Literatur systematisch durchsucht und zusammengefasst (Cooper & Koenka, 2012), um einen Überblick über Korrelate von Lehrer-Schüler-Beziehungen zu schaffen. Hierbei untersuchten wir drei Forschungsfragen

1. Inwieweit hängen akademische, verhaltensbezogene, sozio-emotionale, motivationale und kognitive Schülereigenschaften mit Lehrer-Schüler-Beziehungen in der meta-analytischen Literatur zusammen?

2. Welche Moderatoren beeinflussen diese Zusammenhänge?

3. Welche methodische Qualität haben die einbezogenen Meta-Analysen?

Methode

Um diese Forschungsfragen zu beantworten, analysierten wir 24 Meta-Analysen, die rund 130 Effektstärken für über eine Million Schüler*innen umfassten. Nach der Präregistrierung erfolgte eine systematische Literatursuche. Während mehrerer Runden der Überprüfung mithilfe unserer Ein- und Ausschlusskriterien identifizierten wir 24 passende Meta-Analysen. Aus diesen Meta-Analysen extrahierten wir die Effektstärken zum Zusammenhang von Lehrer-Schüler-Beziehungen und akademische, verhaltensbezogene, sozio-emotionale, motivationale und allgemeine kognitive Schülereigenschaften. Für die Forschungsfragen 1 und 2 haben wir die Ergebnisse zusammengefasst und einen narrativen Überblick erarbeitet. Für Forschungsfrage 3 bewerteten wir die Qualität der Meta-Analysen mit Hilfe der AMSTAR-2 Skala (angepasst an korrelative Studien in der Psychologie und Bildungsforschung; Shea et al., 2017).

Ergebnisse und ihre Bedeutung

Mit Blick auf die Lehrer-Schüler-Beziehungen werden unterschiedliche Ergebnisvariablen analysiert (Forschungsfrage 1). Die stärksten Zusammenhänge zeigten sich für Konflikt und Abhängigkeit in der Lehrer-Schüler-Beziehung mit Verhaltensproblemen der Schüler*innen (r = .35 bis .57; Nurmi, 2012). Positive Lehrer-Schüler-Beziehungen zeigte die stärkste Verbindung mit der Beteiligung in der Schule (r = .26 bis .34; Roorda et al., 2011), prosozialem, externalisierendem und internalisierendem Verhalten (r = .25; Endedijk et al., 2021) sowie mit Lernmotivation in Kombination mit Beteiligung der Schüler*innen (r = .23; Wang et al., 2020). Alter oder Klassenstufe waren die am häufigsten untersuchten Moderatoren mit teilweise gegenläufigen Befunden (Forschungsfrage 2). Geschlechterunterschiede wurden dagegen seltener festgestellt. Gleichzeitig wurde der Effekt der Informationsquelle häufig untersucht, d.h., ob und auf welche Weise Lehrkräfte, Gleichaltrige oder die Schüler*innen selbst die Lehrer-Schüler-Beziehung bewerteten. Für Forschunsgfrage 3 diskutieren wir die Qualitätsunterschiede der Meta-Analysen.

Mit dem systematischen Review von Meta-Analysen fassen wir die Forschungslandschaft zu Korrelaten von Lehrer-Schüler-Beziehungen zusammen und zeigen, in welchem Zusammenhang diese mit Lehrer-Schüler-Beziehungen stehen. Unseren Ergebnissen folgend sollten Lehrkräfte für die Wirkung von Lehrer-Schüler-Beziehungen und deren Zusammenhängen sensibilisiert werden. Einige Interventionen zur Verbesserung von dieser wichtigen Beziehungen wurden bereits meta-analytisch mit vielversprechende Ergebnissen untersucht (Kincade et al., 2020). Ein nächster Schritt ist nun die experimentelle Überprüfung der gefundenen Korrelate, um positive Lehrer-Schüler-Beziehungen als wirksame Strategie zur Verbesserung von akademischen, verhaltensbezogenen, sozio-emotionalen, motivationalen und kognitiven Schülereigenschaften kausal zu bestätigen.

E-19: Bedeutung von Online-Self-Assessments für Studienwahl und Studienerfolg

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* T03 R03 D75

Bedeutung von Online-Self-Assessments für Studienwahl und Studienerfolg

Chair(s): **Karina Karst** (Universität Mannheim, Deutschland), **Stefan Janke** (Universität Mannheim, Deutschland)
*Diskutant*in(nen):* **Martin Neugebauer** (Freie Universität Berlin)

Im deutschsprachigen Hochschulraum setzen Bildungsakteure (u.a. Einzelhochschulen, Ministerien, Agentur für Arbeit) zunehmend auf (Online) Selbstreflexionsverfahren zur Studienorientierung (engl. Online Self-Assessments/OSAs). Dabei soll eine günstige Vorinformation insbesondere hinsichtlich der Passung zwischen eigenen Fähigkeiten und Interessen mit den jeweiligen Studieninhalten eine kohärente und passende Studienwahl sicherstellen. Langfristig sollen korrekte Erwartungen an die Anforderungen des Studiums und Selbstreflexionsprozesse wiederum zu nachhaltigem Studienerfolg und ausbleibendem Studienabbruch beitragen.

Während der Einsatz von OSAs entsprechend durch eine Reihe normativer Annahmen geleitet wird, mangelt es an systematischer, empirischer Forschung zu diesen Annahmen. Im Rahmen des Symposiums soll in vier Beiträgen die Bedeutung von OSAs bei der Studienwahl und für Studienerfolg aus pädagogisch-psychologischer Sicht überprüft werden.

Damit OSAs überhaupt langfristig wirksam sein können, ist es eine zentrale Voraussetzung, dass Studieninteressierte diese Verfahren verwenden und dass die Ergebnisse der Rückmeldung zu ihrer eigenen Passung und mögliche Erwartungsdiskrepanzen in Reflexionsprozesse und in die Studienentscheidung einbezogen werden. Entsprechend sollen die ersten beiden Beiträge des Symposiums die Frage beantworten, welche Faktoren dazu beitragen, dass OSAs in die eigene Entscheidungsfindung einbezogen werden. In Beitrag 1 wird dabei anhand einer groß angelegten Stichprobe von Studieninteressierten analysiert, wovon es abhängt ob Nutzer*innen einen OSA als nützlich einstufen. Aus einer praktischen Perspektive liefert der Beitrag wichtige Informationen dazu, wie sich OSAs adressatengerecht gestalten lassen. In Beitrag 2 wird in einem experimentellen Design die spezifische Wirksamkeit von Feedbackkomponenten in OSAs untersucht (experimentelle Manipulation des Feedbackzeitpunkts). Es zeigt sich, dass die Bereitstellung von Feedback in einigen, aber nicht allen, geprüften Aspekten die Wirksamkeit des OSA erhöht. Diese Erkenntnis ermöglicht eine kritische Einordnung der Bedeutung von Rückmeldungen für die Wirksamkeit von OSAs.

Ob die Bearbeitung eines OSAs tatsächlich eine günstigere Studienwahl herbeiführt und damit in Bewerbungsprozesse eingebunden werden sollte, ist Schwerpunkt von Beitrag 3. Hier zeigt sich anhand der Auswertung von Bewerber*innendaten für einen Masterstudiengang über mehrere Jahre, dass durch die (verpflichtende) Einführung eines OSAs die formale Passung der Studieninteressierten zum Studiengang gesteigert werden konnte. Beitrag 4 geht schließlich der Frage nach, wie intensiv OSAs genutzt werden, wenn sie prozessbezogen als Selbstreflexionsmöglichkeit in den Studienverlauf eingebettet werden. Es zeigte sich dabei, dass das Ausmaß der studentischen Lernzielorientierung in Zusammenhang damit steht, inwieweit Studierende studienbegleitende OSAs für Reflexionsprozesse nutzen. Dies verdeutlicht, dass die optimale, formative Nutzung

entsprechender prozessorientierter OSAs zumindest anteilig auch von der Verarbeitungsmotivation abhängig ist.

Alle vier Beiträge werden am Ende des Symposiums aus soziologischer Perspektive hinsichtlich des gesellschaftlichen Nutzens von OSAs diskutiert.

Beiträge des Symposiums

Wer nutzt und wem nutzen OSA?

Pascale Stephanie Petri, Martin Kersting
Justus-Liebig-Universität Gießen

Die Studienwahl ist die erste Herausforderungen zu Beginn einer akademischen Laufbahn. Ist die Wahl getroffen, müssen sich Studieneinsteiger*innen in der neuen akademischen Umwelt zurechtfinden und den Anforderungen ihres Studienfachs gerecht werden. In keiner anderen Phase des Studiums brechen so viele Studierende ab (Neugebauer, Heublein & Daniel, 2019). Zu den Hauptgründen für Studienabbruch zählen Leistungsdefizite und Motivationsmangel (Heublein et al., 2017). Ursächlich dafür könnte sein, dass Studieneinsteiger*innen von den tatsächlichen Studieninhalten überrascht werden. Daraus können Überforderung und enttäuschte Erwartungen resultieren. Viele Hochschulen versuchen die Studienwahl zu unterstützen und die Diskrepanz zwischen Studienerwartungen und Studienwirklichkeit zu minimieren, indem sie Studieninteressierten die Möglichkeit bieten, sich im Rahmen von Online Self-Assessments (OSA) vorab intensiv mit der Studienwahl auseinander zu setzen (Hasenberg & Schmidt-Atzert, 2014; Stoll et al., 2021). Ob dieses Unterstützungsangebot hilfreich ist, wird in den meisten Evaluationsstudien, die sich auf die Kriteriumsvalidität fokussieren, nicht geprüft (z.B. Petri, Schütte & Beauducel, 2021; Hasenberg, 2012; Zimmerhofer, 2008). Eine Evaluation sollte deutlich breiter aufgestellt sein (Weis, Stoll, Dietrich & Wachtel, 2021; Wittmann, 1990): Beispielsweise kann die Anzahl der Zugriffe auf das Tool und das Nutzer*innen-Urteil als Indikator für die wahrgenommene Nützlichkeit des Angebots dienen.

Auf Basis der Daten der Studierendenbefragungen der Uni Gießen aus den Jahren 2016 bis 2020 (Selbstbericht von insgesamt N = 1776 Studierenden in Fächern, für die ein OSA angeboten wurde) zeigen wir, wie viele der eingeschriebenen Studierenden das OSA vor der Studienwahl genutzt haben. Um der Frage „Wer nutzt OSA?“ nachzugehen, vergleichen wir die Gruppe der Nutzer*innen des OSAs Gießen mit der Gruppe der „Nicht-Nutzer*innen“ desselben Fachs. Hierbei zeigt sich unter anderem, dass der Anteil weiblicher Personen in der Gruppe der OSA-Nutzer*innen höher war und dass die OSA-Nutzer*innen im Vergleich mehr Informationskanäle heranzogen, um sich mit der Studienwahl auseinanderzusetzen. Zusätzlich ziehen wir Daten von über 4000 registrierte Nutzer*innen heran, von denen N = 973 die Möglichkeiten wahrnahmen, detailliertes Feedback zu geben. Zur Klärung der Frage „Wem nutzen OSA?“ werten wir unter anderem die Antworten der Nutzer*innen auf die Fragen a) inwiefern das OSA für sie hilfreich war, um eine genauere Vorstellung von den konkreten Anforderungen und b) ihrer individuellen Passung zu erhalten, aus. Hierbei zeigt sich, dass die Nutzer*innen, die berichten, in dieser Hinsicht besonders vom OSA profitiert zu haben, sich auch im Allgemeinen als eher offen und gewissenhaft beschreiben.

Anhand des vorliegenden Fallbeispiels des OSAs der Uni Gießen resümieren wir, warum eine fest verankerte formative Evaluation, idealerweise unter Einbezug administrativer Daten, wichtig ist, um

die Qualität und die Effekte von OSA umfassend zu untersuchen. Abschließend diskutieren wir, welche Implikationen unsere Erkenntnisse zur Nutzung, zur wahrgenommenen Nützlichkeit und zur Kriteriumsvalidität des OSAs für die Optimierung des Angebotes liefern. Basierend auf bisherigen Erkenntnissen skizzieren wir außerdem eine beispielhafte Evaluationsstrategie mit konkreten Empfehlungen für die zu erhebenden Aspekte.

Der Einfluss von Feedback in Online-Self-Assessments zur Studienwahl auf Einstellungsänderungen bei Studieninteressierten

Belinda Merkle, Hanna Bürkle, Karina Karst, Stefan Janke
Universität Mannheim

Die Entscheidung für das richtige Studienfach ist für viele Studieninteressierte mit großen Unsicherheiten verbunden. Webbasierte Selbstreflexionstests (engl. Online-Self-Assessments, kurz OSAs) sollen bei dieser Entscheidung als Orientierungssysteme dienen und Studieninteressierten helfen, schon vor Studienbeginn ihre Erwartungen an ein Studienfach mit der Studienrealität abzugleichen. Häufig dient dabei eine Befragung zu eigenen Erwartungen an Inhalte und Anforderungen des Studienfaches als Grundlage kriterialer Vergleichsprozesse. Im Anschluss an eine solche Befragung erhalten Studieninteressierte dann ein Feedback, das Informationen über die Passung zwischen ihren Erwartungen und der Studienrealität enthält. Basierend auf der kognitiven Dissonanztheorie (Festinger, 1957) ist anzunehmen, dass besagtes Feedback bei den Studieninteressierten zu Einstellungsänderungen führt, wenn Diskrepanzen zwischen Erwartungen an das Studium und der Studienrealität rückgemeldet werden.

Erste entsprechende Arbeiten konnten diese Annahme stützen und zeigen, dass rückgemeldete Erwartungsdiskrepanzen in der Tat mit Einstellungsänderungen einhergehen (Karst et al., 2017). Unklar ist allerdings, ob es für diese Einstellungsänderung überhaupt eines Feedbacks bedarf oder ob schon die Konfrontation mit unerwarteten Studieninhalten in der vorausgehenden Befragung zu entsprechenden Prozessen führen kann. Um diese Forschungslücke zu schließen, untersuchen wir, in welchem Ausmaß der Einsatz von Feedback den Zusammenhang zwischen den bei der Bearbeitung eines OSAs entstehenden Erwartungsdiskrepanzen und Einstellungsänderungen in Bezug auf das jeweilige Studienfach moderiert. Wir nehmen an, dass der Zusammenhang zwischen Einstellungsänderungen und Erwartungsdiskrepanzen stärker sein sollte, wenn nach der Bearbeitung des OSAs Feedback zum Ausmaß der Erwartungsdiskrepanz gegeben wird.

Wir überprüften unsere Forschungsfragen mit Hilfe eines experimentellen Designs mit 234 Studieninteressierten (M = 19.12 Jahre, 70.9 % weiblich), welche überwiegend die Schule bereits abgeschlossen hatten (51.3 %) oder die 10. - 13. Klasse besuchten. Die Studienteilnehmer:innen wurden randomisiert einer von zwei Gruppen zugeordnet. Beide Gruppen beantworteten zunächst Fragen zu ihrer Einstellung (intrinsischer Wert, Wichtigkeit, Nützlichkeit, Kosten, Erfolgserwartung, Wahlentschlossenheit und Wahlsicherheit) zu einem Studienfach der Wirtschafts- und Sozialwissenschaften. Danach bearbeiteten sie ein OSA. Die Experimentalgruppe (n = 103) erhielt unmittelbar ein Feedback zu ihren Erwartungsdiskrepanzen und beantwortete daraufhin erneut die Fragen zu ihren Einstellungen. Die Kontrollgruppe (n = 131) beantwortete unmittelbar die Fragen zu ihren Einstellungen und erhielt danach das Feedback.

Ausgewertet wurden die Daten mit Hilfe hierarchischer Moderationsanalysen. Dabei zeigen sich Haupteffekte der gemessenen absoluten Erwartungsdiskrepanz (= Betrag der Diskrepanz) auf die spätere Einstellungsänderung (ebenfalls im Betrag): Je höher die absoluten Erwartungsdiskrepanzen, desto größer die absolute Änderung im intrinsischen Wert ($\beta = .21$, $p = .002$), in der persönlichen Wichtigkeit ($\beta = .17$, $p = .010$), in der Nützlichkeit ($\beta = .21$, $p = .002$) und in den Erfolgserwartungen ($\beta = .26$, $p < .001$). Es zeigt sich kein Haupteffekt der Experimentalbedingung. Demnach ändern Studieninteressierte unabhängig von den absoluten Erwartungsdiskrepanzen ihre Einstellung nicht stärker, wenn sie Feedback erhalten (im Vergleich zu keinem Feedback). Es zeigen sich allerdings Moderationseffekte der Experimentalbedingung für den Zusammenhang zwischen den Erwartungsdiskrepanzen mit Wahlsicherheit ($\beta = .18$, $p = .056$) und Erfolgserwartungen ($\beta = .22$, $p = .01611$). Das bedeutet, dass stärkere Erwartungsdiskrepanzen mit einer stärkeren Änderung in der Studienwahlsicherheit und den Erfolgserwartungen einhergehen, wenn die Studieninteressierten Feedback erhalten (vs. kein Feedback erhalten).

Die Ergebnisse legen nahe, dass bereits die Auseinandersetzung mit Studieninhalten Einstellungsänderungen hervorrufen kann. Für andere wenige Einstellungsvariablen hingegen (Nützlichkeit und Erfolgserwartungen) scheint ein Feedback wichtiger, damit sich Erwartungsdiskrepanzen auch in Einstellungsänderungen niederschlagen. Die Befunde verdeutlichen, dass neben der Entwicklung günstiger Feedbackprozeduren, insbesondere die Auswahl inhaltsvalider Items von zentraler Bedeutung für die intendierte Wirksamkeit von OSAs ist.

Anregung zur Selbstselektion und Steigerung der Grundquote – Welchen Beitrag leisten Online Self-Assessments im Kontext der Studierendenauswahl?

Gundula Stoll, Thomas Gfrörer
Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

Online Self-Assessments (OSA) sollen Studieninteressierte zur systematischen Reflexion ihres Studienwunsches anregen (Arnold & Hachmeister, 2004). Durch die Bearbeitung eines OSA sollen sich die Studieninteressierten über ihre Passung zum jeweiligen Studiengang bewusstwerden. Bei schlechter Passung sollte dies zu einem Überdenken der Studienentscheidung führen – bis hin zum möglichen Verzicht auf die Studienbewerbung (Scharlau, Bunte & Wiescholek, 2013; Thiele & Kauffeld, 2020). Es wird angenommen, dass diese durch OSA angeregten Selbstselektionsprozesse die Grundquote (den Anteil potenziell geeigneter Bewerber*innen und Studierender) erhöhen können. Weniger gut geeignete Studieninteressierte sollten sich durch die Bearbeitung des OSA über ihre mangelnde Passung zum Studiengang bewusstwerden und eher auf eine Bewerbung bzw. die Aufnahme des Studiums verzichten; während gut geeignete Studieninteressierte durch die Bearbeitung des OSA in ihrer Bewerbungs- oder Studienabsicht bestärkt werden sollten.

Die Anregung zur Selbstreflexion und Selbstselektion stellt damit eine der wichtigsten Funktionen von OSA dar (vgl. Stoll et al., 2022). Trotzdem fehlen bis heute empirische Belege dafür, dass OSA tatsächlich die intendierten Reflexionsprozesse anstoßen und zur Selbstselektion der Studieninteressierten (bzw. der Steigerung der Grundquote) beitragen.

Wir untersuchen die Selbstselektionsannahme am Beispiel eines zulassungsbeschränkten Masterstudienganges, der sich besonders für diese Untersuchung eignet, weil (1) für diesen

Studiengang klar definierte Zulassungsvoraussetzungen vorliegen, die die Klassifikation potenziell geeigneter Bewerber*innen und somit die Berechnung der Grundquote ermöglicht. (2) Das OSA mit dem Ziel eingeführt wurde, den Anteil der geeigneten Bewerber*innen an den Bewerbungen zu erhöhen, da es zu viele formal nicht geeignete Bewerbungen gab. (3) Für diesen Studiengang Bewerber*innendaten vorliegen, die eine Analyse der Grundquote vor und nach der Einführung des OSA sowie vor und nach der verpflichtenden Verankerung des OSA in der Auswahlsatzung erlauben. Basierend auf Daten aus diesem Studiengang untersuchen wir zwei ineinandergreifende Fragestellungen. Dabei werden sowohl Nutzer*innendaten aus dem OSA als auch unabhängig vom OSA erfasste Bewerber*innendaten aus dem Zulassungs- und Auswahlverfahren herangezogen.

Forschungsfrage 1: Fördert die Bearbeitung des OSA die Selbstreflexion und Selbstselektion? Anhand von Nutzer*innendaten aus dem OSA (N = 703) wird untersucht, ob Personen, die im OSA gut abschneiden, eher in ihrer Studienabsicht bestärkt werden und Personen, die im OSA schlecht abschneiden, angeregt werden, ihre Bewerbungsabsicht nochmal kritisch zu überdenken. Die Ergebnisse zeigen, dass ein substanzieller Anteil (ca. 20%) der Studieninteressierten nach der Bearbeitung des OSA angeben, dass sie das OSA eher davon abgeschreckt hat, sich für den Masterstudiengang zu bewerben. Zudem zeigen sich systematische Zusammenhänge zwischen dem Abschneiden im OSA (als Indikator für die Passung zum Studiengang) und der Bestärkung bzw. Abschreckung hinsichtlich der Bewerbung für den Studiengang.

Forschungsfrage 2: Hat der Einsatz eines OSA als vorgeschaltetes Element im Auswahlverfahren positive Auswirkungen auf die Grundquote der Bewerbungen? Anhand der Bewerber*innendaten (N = 726) vor und nach der Einführung des OSA sowie vor und nach der verpflichtenden Verankerung des OSA in der Auswahlsatzung wird untersucht, inwiefern die Einführung eines OSA mit einer Steigerung der Grundquote einhergeht. Die Ergebnisse zeigen nach Einführung des OSA als freiwilliges Beratungsinstrument einen deutlich höheren Anteil (51%) formal geeigneter Bewerberinnen und Bewerber (Bachelornote ≤ 2,5 und mind. 9 ECTS in Statistik) als vorher (35%) und einen nochmals höheren Anteil (78%) nach der verpflichtenden Verankerung des OSA.

Insgesamt weisen die Befunde der vorliegenden Studie darauf hin, dass der Einsatz eines OSA als vorgeschaltetes Element im Rahmen des Auswahlverfahrens die intendierten Selbstreflexionsprozesse auslösen kann. Zudem zeigt die Analyse der Bewerberdaten durch den Vergleich zwischen Jahren mit und ohne Einsatz des OSA, dass die Einführung des OSA im untersuchten Studiengang tatsächlich mit der erwünschten Steigerung der Grundquote einhergeht. Vor dem Hintergrund dieser Befunde wird diskutiert, wie OSA sinnvoll in Beratungsstrukturen und Bewerbungsprozesse eingebunden werden können und wie das Potential von OSA durch eine optimale Verzahnung mit bestehenden Strukturen verstärkt werden kann.

Online-Self-Assessment als Reflexionstool im Lehramtsstudium

Robert Grassinger
Pädagogische Hochschule Weingarten

Häufig dienen Online-Self-Assessments (OSAs) einer Orientierung für Studieninteressierte. Sie vermitteln vor Studienbeginn Informationen (z.B. zu einem Studiengang) und/oder regen eine Eignungsanalyse an (Hasenberg & Schmidt-Atzert, 2014; Heukamp et al., 2009). Im Gegensatz dazu

begleitet das OSA „PHokus-Self-Assessment“ Lehramtsstudierende während ihres Studiums und zielt darauf ab, mithilfe Selbsteinschätzungen und reflexiver Fragen (Boud & Brew, 1995) selbstgesteuertes Lernverhalten zu unterstützen (Visotschnik et al., in press). Zu drei Zeitpunkten – Beginn des Studiums, im Rahmen des Orientierungspraktikums, beim Übertritt vom Bachelor- in den Masterstudiengang Lehramt – reflektieren Studierende motivationale (z.B. Selbstwirksamkeitserleben, wahrgenommene Nützlichkeit der Lehrinhalte), lernverhaltensbezogene (z.B. Anstrengungsbereitschaft, Zielsetzung und Planung) sowie ressourcenbezogene (z.B. soziale Ressourcen, Stressresistenz) Aspekte selbstgesteuerten Lernverhaltens (Boekaerts, 1999; Schmitz et al., 2007), jeweils spezifisch für die Studiensituation (z.B. Studienbeginn, Praktikum). Diese Idee basiert auf Befunden, wonach OSAs Lernverhalten und Leistungen von Studierenden zu begünstigen vermögen (Ibabe & Jauregizar, 2010; Peat & Franklin, 2002).

Weniger untersucht ist, was beeinflusst, wie intensiv Studierende dieses Angebot zur Reflexion nutzen. Es wird im Beitrag argumentiert, dass lernzielorientiertere Studierende aufgrund ihres Strebens nach Verständnis und Lernzuwachs (Ames, 1992) das OSA stärker zur Reflexion nutzen, arbeitsvermeidungszielorientiertere Studierende hingegen dies qua Definition (Ames, 1992) weniger tun. Die erste Hypothese lautet: Das Ausmaß an Reflexion im Rahmen eines OSAs ist positiv mit einer Lernzielorientierung und negativ mit einer Arbeitsvermeidungszielorientierung assoziiert.

Neben kognitiven Reflexionsprozessen geht die Bearbeitung eines OSAs mit emotionalem Erleben einher und kann Studierende verunsichern. Konkret wird im Beitrag argumentiert, dass die Bearbeitung eines OSAs mit einem Erwartungsabgleich assoziiert sein kann (Karst et al., 2017), der in Anlehnung an Erwartungs-Wert-Modelle der Motivation (Eccles & Wigfield, 2020) sich in nicht erfüllter respektive übertroffener Erfolgserwartung und nicht erfüllter respektive übertroffener Werterwartungen äußert (Grassinger, 2018). Nicht erfüllte motivationale Erwartungen, so die Argumentation, begünstigt das Erleben von Unsicherheit ob des gewählten Studiengangs. Die zweite Hypothese lautet: Nicht erfüllte motivationale Erwartungen stehen in einem positiven Zusammenhang mit dem Erleben von Unsicherheit ob des gewählten Studiengangs.

Methodik

Im Rahmen des OSAs im Kontext des Orientierungspraktikums (N = 165) wurden die Studierenden gebeten, ergänzend zum OSA Items zu ihrer Lernzielorientierung (4 Items), ihrer Arbeitsvermeidungszielorientierung (4 Items) und ihrem erwartungswidrig motivationalen Erleben (je 3 Items) zu beantworten. Nach Bearbeitung des OSAs bekamen die Studierenden Fragen zum Ausmaß der Reflexion (3 Items, Beispielitem: „Durch die Fragen habe ich gezielt darüber nachgedacht, was ein erfolgreiches Studium begünstigen kann“) und der Verunsicherung (3 Items, Beispielitem: „Die Fragen haben mich hinsichtlich meiner Studienwahl verunsichert“). Die Daten wurden mit einem Strukturgleichungsmodell analysiert. Hierbei wurde das Ausmaß an Reflexion regressiert auf Lernzielorientierung und Arbeitsvermeidungszielorientierung und das Erleben von Verunsicherung auf erwartungswidriges Erleben der Erfolgserwartung und der Werterwartungen. Letztgenanntes erwartungswidriges motivationales Erleben wurde zudem mit dem Ausmaß an Reflexion korreliert.

Ergebnisse

Es zeigte sich, dass lernzielorientiertere Studierende mehr und arbeitsvermeidungszielorientiertere Studierende weniger Reflexion berichteten. Zudem berichteten die Studierenden eine umso größere Verunsicherung ob ihrer Studienwahl, je stärker die Erfolgserwartung nicht erfüllt wurde und je mehr Arbeitsvermeidungsziele sie verfolgten. Weiterhin zeigte sich, dass weder erwartungswidriges Erleben der Werterwartungen noch das Ausmaß an Reflexion mit dem Erleben von Verunsicherung assoziiert waren. Das Ausmaß an Reflexion stand zudem positiv mit erwartungswidrigem Erleben der Werterwartungen in Zusammenhang, nicht jedoch mit erwartungswidrigem Erleben der Erfolgserwartung.

Diskussion

Das Angebot mithilfe eines OSAs die Selbststeuerung des eigenen Lernverhaltens im Studium zu reflektieren scheint vorwiegend von Studierenden mit hoher Lernzielorientierung und weniger von Studierenden mit hoher Arbeitsvermeidungszielorientierung genutzt. Folglich mag es wertvoll sein, in der Instruktion zu einem OSA Lernziele salient zu machen. Weiterhin verweisen die Befunde darauf, dass OSAs als Tool zur Unterstützung der Selbststeuerung auch ob der Wahl des Lehramtsstudiengangs verunsichern können und dies vor allem dann tun, wenn dadurch nicht erfüllte motivationale Erwartungen salient werden.

E-20: Wohlbefinden und Emotionen von Kindern und Jugendlichen in Zeiten der Corona-Pandemie

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* R11 T00 D05

Wohlbefinden und Emotionen von Kindern und Jugendlichen in Zeiten der Corona-Pandemie

Chair(s): **Linda Wirthwein** (TU Dortmund, Germany), **Ricarda Steinmayr** (TU Dortmund, Germany)
*Diskutant*in(nen):* **Thamar Voss** (Universität Freiburg)

Im Kontext internationaler Schulleistungsvergleichsstudien gewinnt das subjektive Wohlbefinden (SWB) von Schülerinnen und Schülern (SuS) vermehrt an Aufmerksamkeit (OECD, 2017). So liegt der Fokus nicht nur darauf, dass SuS optimale Lern- und Leistungsergebnisse erzielen, sondern sich auch in der Schule wohl fühlen, zumal ein hohes Wohlbefinden eine zentrale Voraussetzung für erfolgreiches Lernen darstellt (Krapp, 2005). Erste Befunde weisen auf eine Abnahme des Wohlbefindens von SuS im Verlauf der Corona-Pandemie hin. In diesem Kontext stellt sich insbesondere die Frage, welche Variablen Einfluss auf diese Abnahme haben. Hier setzen die vier Beiträge des vorliegenden Symposiums an. Studie 1 („Subjektives Wohlbefinden von Jugendlichen im Verlauf der Corona-Pandemie“) betrachtet das SWB von *N* = 266 Jugendlichen vor und zwei Jahre nach Beginn der Pandemie in einem Längsschnitt. Darüber hinaus werden Persönlichkeitsvariablen und der soziodemographische Hintergrund untersucht, die Einfluss auf diese Abnahme haben könnten. Beitrag 2 („Bedeutung zentraler Ressourcen für den Schulerfolg von Lernenden in herausfordernden Zeiten“) betrachtet an einer Stichprobe von *N* = 220 Jugendlichen neben personalen (u.a. Selbstwirksamkeit) auch soziale Ressourcen (u.a. Unterstützung durch Freunde), die für die Wahrnehmung der Corona-Pandemie als bedrohlich relevant sein könnten. Zusätzlich wird analysiert, inwiefern diese Ressourcen mit dem SWB und Lernerfolg assoziiert sind. Studie 3 („Sozio-emotionale Entwicklungsprozesse im Grundschulalter infolge der ersten Schulschließung während der COVID-19 Pandemie“) untersucht an *N* = 590 Grundschulkindern weitere sozio-emotionale Variablen (u.a. Optimismus, Beziehung zu Erwachsenen) und überprüft, welchen Einfluss der sozioökonomische Status auf die Wahrnehmung dieser Ressourcen hat. Der letzte Beitrag („Zurück im Klassenzimmer: Zusammenhänge von Lehrpersonenemotionen mit Unterrichtsverhalten und dem Wohlbefinden der Schüler*innen in der Primarstufe“) nimmt die Rolle von Lehrpersonenemotionen in den Blick und deren Zusammenhänge mit u.a. positiven Emotionen der SuS. Die Stichprobe umfasst *N* = 134 Lehrkräfte und *N* = 1550 Kinder.

Die Beiträge des Symposiums beziehen sich dabei sowohl auf das emotionale Erleben von Jugendlichen als auch von Grundschulkindern. Darüber hinaus werden zusätzlich Lehrkraftemotionen und ihre Auswirkungen auf Emotionen der Kinder fokussiert. Während sich die ersten beiden Beiträge auf eine hedonistische Definition von Wohlbefinden beziehen – operationalisiert als Zufriedenheit und Stimmung – greifen Studie 3 und 4 auf eine eudaimonische Definition zurück und untersuchen u.a. Optimismus und intrinsische Motivation. Die Beiträge widmen sich außerdem moderierenden Faktoren im Erleben der SuS und untersuchen sowohl personale (z.B. Persönlichkeitseigenschaften), soziale (z.B. Unterstützung durch Familie und Freunde) als auch soziodemographische Variablen. Anhand der Ergebnisse kann auf relevante Faktoren hingewiesen werden, welche SuS eine zusätzliche Unterstützung benötigen, um die emotionalen Folgen der Corona-Pandemie aufzufangen.

Subjektives Wohlbefinden von Jugendlichen im Verlauf der Corona-Pandemie – Ist nach zwei Jahren alles wieder gut?

Linda Wirthwein, Patrick Paschke, Ricarda Steinmayr
TU Dortmund

Befunde weisen auf eine Abnahme des subjektiven Wohlbefindens und der mentalen Gesundheit von Kindern und Jugendlichen nach Beginn der SARS-CoV-2-Pandemie und den damit einhergehenden Infektionsschutzmaßnahmen (u.a. Schulschließungen, Kontaktbeschränkungen) hin (Shoshani & Kor, 2021; Steinmayr et al., 2022). Immer noch liegen jedoch vergleichsweise wenige Studien vor, die das subjektive Wohlbefinden (SWB) von Kindern und Jugendlichen vor Beginn und im Verlauf der Pandemie im Längsschnitt untersuchen. Bislang wurde das Absinken des Wohlbefindens vor allem auf einen reduzierten sozialen Austausch zurückgeführt, da soziale Beziehungen wichtige Determinanten des SWBs im Kindes- und Jugendalter darstellen (vgl. Konu et al., 1999). Aktuell ergibt sich jedoch vor allem die Frage, ob das SWB von Schülerinnen und Schülern (SuS) nach Ende der Kontaktbeschränkungen und Schulschließungen immer noch niedriger ausgeprägt ist als vor dem Beginn der Pandemie oder ob die SuS ähnlich glücklich sind wie vor dem Beginn der Pandemie. Im Sinne der „Set Point Theory“, welche annimmt, dass das SWB zwar prinzipiell relativ stabil ist, sich aber bestimmte Lebensereignisse (wie eine weltweite Pandemie) langfristig negativ auf das SWB auswirken können (Diener et al., 2006; Luhmann et al., 2012), ist zu vermuten, dass das SWB nach zwei Jahren Pandemie immer noch nicht den Ausgangswert von vor der Pandemie erreicht hat.

Die Abnahme des SWBs scheint jedoch auch von der Ausprägung relevanter Moderatorvariablen (z.B. Persönlichkeitsvariablen) abzuhängen (Romm et al., 2021). Etwaige Moderatoren wurden bislang ebenfalls selten im Rahmen eines längsschnittlichen Designs an SuS untersucht. Ziel der vorliegenden Studie ist es daher, das Wohlbefinden von SuS vor dem Beginn und im Verlauf der Corona-Pandemie zu vergleichen und relevante Moderatorvariablen (Persönlichkeit, soziodemographischer Hintergrund) in die Analyse miteinzubeziehen.

Bezugnehmend auf die „Set Point Theory“ nehmen wir an, dass das SWB der SuS im späteren Verlauf der Pandemie immer noch geringer ausgeprägt ist als vor Beginn der Pandemie. Darüber hinaus vermuten wir einen Einfluss von Persönlichkeitsvariablen auf die Veränderung des SWBs.

Wir befragten N = 266 SuS (148 weiblich, 114 männlich, 4 gaben ihr Geschlecht nicht an) zweier Gymnasien in NRW zu zwei Messzeitpunkten. Zum ersten Messzeitpunkt (t1; Oktober 2019) waren die SuS in Stufe 10 und zum zweiten Messzeitpunkt (t2; Januar/Februar 2022) in Stufe 12. Das mittlere Alter zu t1 betrug 15.16 Jahre (SD = 0.534). Zu beiden Messzeitpunkten wurden Daten zum Stimmungsniveau und zur allgemeinen Lebenszufriedenheit mit einer gekürzten Form der habituellen Wohlbefindensskala von Dalbert (1992) erhoben. Soziodemographische Angaben (Alter, Geschlecht, Anzahl der Bücher zu Hause, höchster Schulabschluss der Eltern) wurden ebenfalls von den Jugendlichen erbeten.

Die Ergebnisse der latent change score models (LCMs) zeigten, dass sowohl die Lebenszufriedenheit ($\mu\Delta$ = -.18; p = .043) als auch das Stimmungsniveau ($\mu\Delta$ = -.26; p = .003) der SuS von t1 zu t2 signifikant abgenommen haben. Darüber hinaus zeigten die Regressionsanalysen, dass t1 Extraversion über die anderen Prädiktoren hinweg (t1 Lebenszufriedenheit, t1 Neurotizismus,

Schulabschluss der Eltern, Anzahl Bücher im Haushalt, Geschlecht) sowohl die t2 Lebenszufriedenheit (β = .22; p = .010) als auch das t2 Stimmungsniveau (β = .26; p = .010) positiv vorhersagte. Zudem sagte ein höherer Schulabschluss der Eltern die Lebenszufriedenheit, nicht aber das Stimmungsniveau der SuS negativ vorher (β = -.14; p = .028). Die übrigen Prädiktoren hatten keine signifikante Vorhersagekraft für die Lebenszufriedenheit oder das Stimmungsniveau.

Diskutiert werden die Befunde u.a. vor dem Hintergrund der „Set Point Theory“. In diesem Zusammenhang stellt sich insbesondere die Frage, für wen sich die SARS-CoV-2-Pandemie weiterhin langfristig negativ auf das SWB auswirken wird und wie diese Kinder und Jugendlichen besonders unterstützt werden könnten.

Bedeutung zentraler Ressourcen für den Schulerfolg von Lernenden in herausfordernden Zeiten

Justine Stang-Rabrig[1], Sebastian Nicolas Thomas Vogel[1], Marco Forciniti[2], Nele McElvany[1]
[1]TU Dortmund, [2]Ursulinengymnasium, Werl

Der Schulerfolg ist ein multidimensionales Konstrukt, das leistungsbezogene (z.B. Lernaktivitäten und erfolg), aber auch affektiv-motivationale (z.B. Wohlbefinden) Outcomes umfasst, die wiederum zentrale Bildungsziele darstellen (z.B. OECD, 2017). Das Wohlbefinden beschreibt, wie eine Person über das Leben allgemein oder in Bezug auf einen bestimmten Bereich denkt und fühlt. Entsprechend stellt die Lebenszufriedenheit einen bedeutsamen Aspekt des Wohlbefindens dar (z.B. Diener et al., 1999). Lernaktivitäten beziehen sich auf die Häufigkeit der Durchführung qualitätsvoller Aktivitäten wie etwa dem Üben und Anwenden von Gelerntem, die der Festigung von Lerninhalten dienen (z.B. Conole, 2007). Der Lernerfolg bildet ab, ob und wie viel Lernende in einem bestimmten Zeitraum gelernt haben und ist somit ein wichtiger Indikator der Schulleistung (z.B. York et al., 2015).

Die drei Schüleroutcomes, die neben anderen Faktoren den Schulerfolg konstituieren, sind insbesondere auch in herausfordernden Zeiten wie der COVID-19-Pandemie von Bedeutung, in denen Lernen vornehmlich nicht in der gewohnten Umgebung, den Klassenzimmern, stattfand. Herausforderungen wie die Pandemie beziehungsweise die Wahrnehmung dieser als belastend können dabei die verschiedenen Outcomes negativ beeinflussen (z.B. Ludewig et al., 2022; Ravens-Sieberer et al., 2021). Daher ist es essenziell, Faktoren zu identifizieren, die positiv mit den Schüleroutcomes zusammenhängen und somit protektiv wirken können.

Wichtige benevolente Faktoren, die vor allem in belastenden Zeiten von Relevanz sind, stellen personale (z.B. Resilienz, Selbstwirksamkeit) und soziale Ressourcen (z.B. Unterstützung durch Familie/Freunde/Lehrkräfte/signifikante Andere) dar. Theoretische Modelle und empirische Studien legen nahe, dass Ressourcen einerseits mit der Wahrnehmung herausfordernder Situationen und andererseits mit zentralen Outcomes wie beispielsweise dem Wohlbefinden und der Leistung zusammenhängen (z.B. Hobfoll, 1989, 2002; Zysberg & Schwabsky, 2021).

Vor diesem Hintergrund wurden folgende Forschungsfragen untersucht: 1) Wie hängt die Wahrnehmung der COVID-19-Pandemie mit Wohlbefinden, Lernaktivitäten und Lernerfolg zusammen? 2) Inwiefern besteht ein Zusammenhang zwischen personalen (Resilienz, Selbstwirksamkeit) beziehungsweise sozialen (Unterstützung durch Familie/Freunde/Lehrkräfte/signifikante Andere) Ressourcen und der Wahrnehmung der COVID-19-

Pandemie? 3) Wie sind die Ressourcen (Resilienz, Selbstwirksamkeit, Unterstützung durch Familie/Freunde/Lehrkräfte/signifikante Andere) mit dem Wohlbefinden, Lernaktivitäten und Lernerfolg assoziiert und wird der Zusammenhang durch die Wahrnehmung der COVID-19-Pandemie mediiert?

Datengrundlage bildeten Angaben von 220 Schüler*innen (60.9% weiblich, 1.4% divers), die im Mittel 16.21 Jahre alt waren (SD = 0.88). Die Lernenden füllten im Rahmen einer Schulstunde einen Online-Fragebogen aus und machten dabei Angaben zu soziodemografischen Daten und den interessierenden Variablen. Die Items wurden auf vier- beziehungsweise fünfstufigen Likert-Skalen gemessen (1 = überhaupt nicht zufrieden/trifft überhaupt nicht zu/nie bis 4 bzw. 5 = voll und ganz zufrieden/trifft voll und ganz zu/immer). Die Reliabilitäten lagen in einem zufriedenstellenden bis guten Bereich (.71 ≤ α ≤ 91). Zur Beantwortung der Forschungsfragen wurden manifeste Strukturgleichungsmodelle in Mplus 8.1 (Muthén & Muthén, 1998–2017) spezifiziert. Die Modellfits waren zufriedenstellend.

Die Strukturgleichungsmodelle zeigten, dass die Wahrnehmung der COVID-19-Pandemie als belastend statistisch signifikant negativ mit dem Wohlbefinden zusammenhing (ß =-.27, $p < .001$; R^2 = .14). Die Wahrnehmung der COVID-19-Pandemie als belastend wurde wiederum durch die personale Ressource Selbstwirksamkeit negativ (ß =-.22, $p < .01$), durch die Unterstützung durch signifikante Andere hingegen positiv vorhergesagt (ß = .32, $p < .001$; R^2 = .22). Sowohl personale als auch soziale Ressourcen (Resilienz, Selbstwirksamkeit, Unterstützung durch Familie/Lehrkräfte/Andere) standen in einem bedeutsamen positiven Zusammenhang mit den Schüleroutcomes Wohlbefinden (R^2 = .57), Lernaktivitäten (R^2 = .25) und Lernerfolg (R^2 = .16), wobei nur die Selbstwirksamkeit mit allen drei Outcomes assoziiert war (ß ≥. 13, $p \leq .05$). Die Wahrnehmung der COVID-19-Pandemie vermittelte Zusammenhänge zwischen Selbstkonzept beziehungsweise Unterstützung durch Andere und dem Wohlbefinden.

Zusammenfassend sind für die erfolgreiche Bewältigung außergewöhnlicher Situationen wie der COVID-19-Pandemie persönliche Ressourcen wie Resilienz und Selbstwirksamkeit, aber auch soziale Ressourcen wie die Unterstützung durch Familie und Lehrkräfte wichtig. Die Ergebnisse werden auf methodischer und inhaltlicher Ebene diskutiert, Implikationen für Forschung und Praxis vorgestellt.

Sozio-emotionale Entwicklungsprozesse im Grundschulalter infolge der ersten Schulschließung während der COVID 19 Pandemie

Christine Schmid, Nanine Lilla, Helvi Koch, Marianne Schüpbach
Freie Universität Berlin

Unsichere Zeiten bilden eine Herausforderung für die sozio-emotionale Entwicklung von Menschen. Manche Menschen wachsen an Krisen, andere gehen geschwächt aus ihnen hervor. Wie gut Krisen bewältigt werden, hängt jeweils von der persönlichen Widerstandskraft und den Ressourcen ab, die mobilisiert werden können. Diese sind jedoch gesellschaftlich nicht gleichverteilt.

Im Rahmen des vorliegenden Beitrags werden vier Skalen genutzt, die aus einem Instrument zur ganzheitlichen Erfassung entwicklungsbezogener sozio-emotionaler Ressourcen und Risikofaktoren bei Schüler*innen stammen (Holistic Student Assessment, HSA; Noam, Allen & Triggs, 2018): Bewegungsorientierung, Beziehung zu Erwachsenen, Optimismus und Ausdauer. Die Auswahl der

Skalen erfolgte vor dem Hintergrund ihrer Zuordnung zu den Dimensionen des zugrundeliegenden „Clover Model“ (Noam & Triggs, 2018; Malti & Noam, 2009). Pro Dimension wurde eine Skala gewählt. Das Clover Model geht von vier basalen Notwendigkeiten oder Bedürfnissen der menschlichen Entwicklung aus, die von den Autoren als Aktivität, Zugehörigkeit, Reflexion und Durchsetzungsvermögen beschrieben wurden.

Das HSA kann einerseits genutzt werden, um im schulischen Kontext Stärken und Schwächen bzw. den Unterstützungsbedarf einzelner Schüler*innen zu diagnostizieren. Andererseits bildet es ein Instrument, das großflächig eingesetzt Bereiche identifizieren hilft, die im Rahmen des sozio-emotionalen Lernens (SEL) im Unterricht sowie auch in außerunterrichtlichen Settings besonderer Förderung bedürfen (Malti & Noam, 2016; Noam, Allen & Triggs, 2018). Die verwendeten Skalen weisen gemäß Validierungsstudien Zusammenhänge mit problematischen internalisierenden und externalisierenden Verhaltensweisen wie beispielsweise Ängstlichkeit, Depression, Delinquenz und Schulabstinenz auf (Malti, Zuffianó & Noam, 2017; Noam, Malti & Guhn, 2012). Sie können daher, je nach Ausprägung, auch als Schutz- oder Risikofaktoren für die weitere kindliche Entwicklung interpretiert werden.

Fragestellung:

In der vorliegenden Studie wird mittels eines Pre-Post-Test-Designs (Little et al., 2020) zunächst untersucht, inwieweit die Schüler*innen selbst eine Verbesserung oder Verschlechterung ihrer sozio-emotionalen Ressourcen infolge der ersten Schulschließung im Frühjahr/Sommer 2020 wahrnahmen. In einem zweiten Schritt wird dann überprüft, welchen Einfluss der sozioökonomische Status der Schüler*innen auf diese Wahrnehmungen hat und ob sich Hinweise für eine Verschärfung der Lage für die sozioökonomisch schlechter gestellten Schüler*innen im Vergleich zu den besser gestellten durch die Pandemie ergeben.

Methode:

Die Analysen basieren auf einer Stichprobe von N=590 Grundschüler*innen (Klassenstufe 4 bis 6) aus Berlin und Brandenburg. Die Daten wurden in sechs Schulen in insgesamt 36 Klassen mittels Paper-and-Pencil-Fragebögen erhoben. Mithilfe von Mplus wurden Latente Change-Score Modelle (Geiser, 2010) berechnet. Fehlende Werte wurden mittels FIML geschätzt, außerdem wurde der geclusterten Datenstruktur Rechnung getragen. Der sozioökonomische Status wurde mittels eines Instruments erhoben, das bereits in der Studie EduCare (Schüpbach, 2010; von Allmen, 2017) zum Einsatz kam.

Ergebnisse:

Für die Gesamtgruppe der befragten Schüler*innen ergab sich intraindividuell eine signifikante Abnahme des Optimismus sowie eine leichte Zunahme der Ausdauer. Interindividuelle Unterschiede in Abhängigkeit von SES traten bei der Bewegungsorientierung und der Beziehung zu Erwachsenen auf: je höher der SES, desto mehr Bewegungsorientierung und desto besser die berichteten Beziehungen zu Erwachsenen. Ein Effekt des SES auf die Veränderungen ergab sich tendenziell bei der Bewegungsorientierung und signifikant beim Optimismus: In beiden Fällen weist ein negatives Vorzeichen auf eine etwas stärkere Abnahme bei höherem SES hin.

Die Ergebnisse sind dahingehend zu interpretieren, dass die Pandemie offensichtlich einen dämpfenden Effekt auf den Optimismus der Schüler*innen hatte. Auf die Ausdauer hingegen wirkte sie sich tendenziell stärkend aus. Außerdem hatten die Schüler*innen mit niedrigerem SES schon vor der Pandemie eine schlechtere Ausgangslage: sie berichteten über eine geringere Bewegungsorientierung sowie über weniger gute Beziehungen zu Erwachsenen als Schüler*innen mit höherem SES. Schließlich erwiesen sich die Schüler*innen mit höherem SES als etwas vulnerabler in ihrem Optimismus sowie tendenziell auch in ihrer Bewegungsorientierung. Im Falle des Optimismus indizieren die Ergebnisse eine Vergrößerung der Unterschiede nach SES bei gleichem Ausgangsniveau, im Falle der Bewegungsorientierung dagegen eine Verringerung der Unterschiede bei deutlich unterschiedlichem Ausgangsniveau.

Zurück im Klassenzimmer: Zusammenhänge von Lehrpersonenemotionen mit Unterrichtsverhalten und dem Wohlbefinden der Schüler*innen in der Primarstufe

Julia Holzer, Luisa Grützmacher, Manfred Prenzel, Barbara Schober, Marko Lüftenegger
Universität Wien

Theoretischer Hintergrund

COVID-19 und die damit verbundenen Einschränkungen stellten Schulen und Unterricht vor besondere Herausforderungen. Besonders ungünstig wirkte sich die Pandemie auf Lernende in herausfordernden Lagen aus. So gaben etwa nur 7% österreichischer Lehrpersonen in einer Befragung an, zuversichtlich gewesen zu sein, bildungsbenachteiligten Schüler*innen den relevanten Stoff gut vermittelt zu haben (Holzer & Spiel, 2020). Nun, nach der Rückkehr der Schüler*innen ins Klassenzimmer, erscheint es umso relevanter, den Blick auf Schulen in herausfordernden Lagen zu richten und Determinanten von Unterrichtsqualität und Schüler*innenwohlbefinden zu untersuchen. Die vorliegende Studie nimmt dafür die Rolle von Lehrpersonenemotionen in den Blick. Diese gelten als wichtige Einflussfaktoren für Unterrichtsverhalten, sowohl hinsichtlich kognitiver und motivationaler Stimulation, als auch Klassenmanagement und sozialer Unterstützung. Positive Emotionen ermöglichen Lehrpersonen den effektiven Einsatz vielfältiger Unterrichtsstrategien um flexibel auf unerwartete Ereignisse zu reagieren und auf die Bedürfnisse der Lernenden einzugehen. Negative Emotionen gelten diesbezüglich als hinderlich (Frenzel et al., 2008). Dementsprechend belegen eine Bandbreite empirischer Studien Zusammenhänge zwischen Lehrpersonenemotionen und Unterrichtsverhalten (Frenzel, Goetz, Stephens, et al., 2009; Kunter et al., 2008). Über die Zusammenhänge von Lehrpersonenemotionen mit Unterrichtsverhalten hinaus, schlagen Frenzel et al. (2008) in ihrem Modell vor, dass Lehrpersonenemotionen sich sowohl direkt durch emotionale Ansteckung, als auch indirekt, durch das Unterrichtsverhalten, auf die Motivation und das emotionale Erleben der Schüler*innen auswirken. Damit beziehen sich Frenzel et al. (2008) auf Schüler*innenmerkmale, die als Komponenten eudaimonischen und hedonischen Wohlbefindens von Schüler*innen verstanden werden (Ryan & Deci, 2001). Während empirische Belege für den Aspekt der emotionalen Ansteckung vereinzelt vorliegen (Frenzel, Goetz, Lüdtke, et al., 2009), liegen keine Studien vor, die Lehrpersonenemotionen, Unterrichtsverhalten und Schüler*innenwohlbefinden im Sinne intrinsischer Motivation und positiver Emotionen kombiniert hinsichtlich direkter, als auch indirekter Effekte untersuchen.

Fragestellung

Entsprechend theoretischer Annahmen (Frenzel et al., 2008) prüft die vorliegende Studie Zusammenhänge von Lehrpersonenemotionen (Freude, Ärger, Angst) mit dem Unterrichtsverhalten (kognitive und motivationale Stimulation, Klassenmanagement, soziale Unterstützung) und dem Schüler*innenwohlbefinden (positive Emotionen und intrinsische Motivation) in einer Stichprobe von Lehrpersonen und Schüler*innen aus Schulen in herausfordernden Lagen. Wir nehmen sowohl direkte Effekte von Lehrpersonenemotionen auf Unterrichtsverhalten und Schüler*innenwohlbefinden an, als auch indirekte Effekte von Lehrpersonenemotionen auf Schüler*innenwohlbefinden, vermittelt über das Unterrichtsverhalten.

Methode

134 Lehrpersonen (93.1% weiblich, 6.9% männlich; 22 – 62 Jahre, Mage = 40.31, SDage = 12.05) und 1550 Schüler*innen (48.3% weiblich, 51.7% männlich; 7 – 12 Jahre, Mage = 9.59, SDage = .90) aus 134 Klassen in österreichischen Volksschulen in herausfordernden Lagen wurden im Kontext eines größeren Forschungsprojekts in Kooperation mit dem Bundesministerium für Bildung, Wissenschaft und Forschung mittels Online-Fragbögen befragt. Die angenommenen Effekte wurden mittels Mehrebenen-Mediationsanalyse untersucht.

Ergebnisse und ihre Bedeutung

Hypothesenkonform identifizierten wir einige direkte Effekte von Lehrpersonenemotionen auf Unterrichtsverhalten. Entsprechend zeigte das Modell moderate bis große Effekte von Freude auf Aspekte kognitiver und motivationaler Stimulation, Klassenmanagement und sozialer Unterstützung. Zudem identifizierten wir moderate negative Effekte von Ärger und Angst auf das Klassenmanagement. Entgegen unserer Annahmen identifizierten wir keine signifikanten Effekte des Unterrichtsverhaltens auf das Schüler*innenwohlbefinden. Hinsichtlich direkter Effekte von Lehrpersonenemotionen auf Schüler*innenwohlbefinden zeigte sich überraschenderweise ein positiver direkter Effekt von Ärger auf intrinsische Motivation der Schüler*innen. Entgegen unserer Annahmen wurden keine indirekten Effekte von Lehrpersonenemotionen auf das Schüler*innenwohlbefinden, vermittelt über das Unterrichtsverhalten, identifiziert.

Hinsichtlich der Evidenz für Zusammenhänge zwischen Lehrpersonenemotionen und Unterrichtsverhalten reiht sich die Studie in den bisherigen Stand der Forschung ein (Frenzel, Goetz, Stephens, et al., 2009; Kunter et al., 2008). Überraschend ist der Befund, wonach Ärger mit intrinsischer Motivation der Schüler*innen positiv zusammenhängt. Ein möglicher Erklärungsansatz könnte im Konzept der emotionalen Ansteckung liegen, wobei im Fall der vorliegenden Studie nicht die negative Valenz der Emotion, jedoch aber der Grad des Arousals der Lehrpersonen auf die Schüler*innen übertragen wird (siehe z.B. Prochazkova & Kret, 2017). Weitere Erklärungsansätze werden diskutiert.

Sonderveranstaltung

E-21: Sonderveranstaltung: Esther Wojcicki

Zeit: Mittwoch, 01.03.2023: 15:30 - 17:15 · *Ort:* R14 R00 A04 & **Hybrid (ZOOM)**

How are we going to work with ChatGPT in the education world?

Esther Wojcicki
Palo Alto High School, United States of America

How are we going to work with ChatGPT in the education world? Is banning it the answer? How effective was it to ban alcohol in the 20th century?

Esther Wojcicki will discuss how education needs to incorporate TRICK (her acronym for success) into the classroom and into parenting to promote thinkers to solve our 21st century challenges.

TRICK means to be trusted and respected, given independence, collaborated with instead of dictated to and treated with kindness. She will discuss ways of incorporating TRICK into the classroom and deal with the challenge of ChatGPT.

Session F

Einzelbeiträge

F-01: Zuwanderung und kulturelle Diversität

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S06 S00 B29
Chair der Sitzung: Beyhan Ertanir

Das Zusammenspiel von kulturellen Stressoren bei Familien mit Zuwanderungshintergrund und ihre Auswirkungen auf die mentale Gesundheit

Beyhan Ertanir[1], Alan Meca[2], Cory Cobb[3], Ingrid Zeledon[4], Pablo Montero-Zamora[5], Jennifer Unger[4], Daniel Soto[4], Byron L. Zamboanga[6], Dilan Aksoy[1], Wassilis Kassis[1], Elma Lorenzo-Blanco[5], Maria Duque[5], Maria Fernanda Garcia[7], Seth Schwartz[5]

[1]Pädagogische Hochschule Fachhochschule Nordwestschweiz; [2]The University of Texas at San Antonio; [3]Auburn University; [4]University of Southern California; [5]The University of Texas at Austin; [6]University of Arkansas; [7]University of Miami

Sowohl die *Cultural Stress Theory* (Salas-Wright & Schwartz, 2019) als auch das *Family Stress Modell* (Conger et al., 2010) postulieren, dass sich (kulturelle) Stressoren negativ auf die mentale Gesundheit von Erwachsenen sowie Jugendlichen mit Zuwanderungshintergrund auswirken kann und dass dieser Zusammenhang über dysfunktionale Familienbeziehungen vermittelt wird. In der Tat belegen eine Vielzahl an empirischen Studien (Cano et al., 2015; Lorenzo-Blanco et al., 2016; 2017; Schwartz et al., 2015; Yip et al., 2018), dass kulturelle Stressoren (z.B. wahrgenommene Diskriminierung, negativer Aufnahmekontext und bikulturelle Stressoren) einen Risikofaktor für die eigene mentale Gesundheit von Menschen mit Zuwanderungshintergrund darstellt. Bislang ist allerdings noch unklar, (a) welche direktionalen Zusammenhänge zwischen den einzelnen kulturellen Stressoren auf intra- und interindividueller Ebene vorliegen und (b) wie die eigenen kulturellen Stressoren mit den kulturellen Stressoren anderer Familienmitglieder zusammenhängen. Des Weiteren ist bis jetzt noch nicht erforscht, ob sich (c) individuelle kulturelle Stressoren auf die mentale Gesundheit anderer Familienmitglieder auswirken. Der vorliegende Beitrag adressiert diese Forschungslücken und zieht für die Beantwortung der Fragestellungen die Daten der amerikanischen Längsschnittstudie COPAL (Contruyendo Oportunidades Para los Adolescentes Latinos; Schwartz et al., 2015; Laufzeit 2010-2013) heran, die Akkulturationsprozesse bei kürzlich eingewanderten hispanischen Familien (N = 302 Eltern-Kind-Dyaden) mit Zuwanderungshintergrund untersucht. Hierfür wurden sowohl Jugendliche (53 % Jungen; M = 14.51, SD = .88) als auch ihre Eltern (73 % Frauen; M = 41.09, SD = 7.13) zu ihrem kulturellen Stressempfinden sowie ihrer mentalen Gesundheit (z.B. depressive Symptomatiken, Selbstwertgefühl, Optimismus, Substanzkonsum, externalisierendes Verhalten, Lebenszufriedenheit) über fünf Messzeitpunkte (im Abstand von sechs Monaten) hinweg befragt (für eine ausführliche Darstellung der Instrumente siehe Schwartz et al., 2015). Die Ergebnisse von Random-Intercept-Cross-Lagged Modelle deuten auf ein komplexes Zusammenhangsmuster hin: Für die Jugendlichen zeigten sich sowohl auf intra- als auch auf interindividueller Ebene signifikante Zusammenhänge zwischen den einzelnen Stressoren (z.B. bidirektionale kreuzverzögerte Effekte zwischen wahrgenommener Diskriminierung und bikulturellen Stressoren sowie negativem Aufnahmekontext), während sich für die Eltern lediglich interindividuelle Zusammenhänge beobachten liessen. Zudem gab es einen marginal signifikanten Cross-Reporter-Kreuzeffekt zwischen von dem vom Jugendlichen berichteten negativen Aufnahmekontext auf die elterliche wahrgenommene Diskriminierung. Des Weiteren zeigte sich, dass kulturelle Stressoren nicht nur einen Effekt auf die eigene mentale Gesundheit haben, sondern auch spezifische

psychosoziale Outcomes anderer Familienmitglieder prädizieren (z.B. sagte der akkulturative Stress von Eltern den Substanzkonsum von Jugendlichen vorher). Im Rahmen des Beitrags sollen die Ergebnisse sowie ihre Implikationen für die Praxis und Forschung diskutiert werden.

Hohe Bildungsaspirationen und optimistische Bildungsentscheidungen – eine Ressource für Familien mit Zuwanderungsgeschichte?

Katharina Kohl, Lisa Benckwitz, Olaf Köller
IPN Kiel, Germany

Empirische Studien zeigen recht konsistent, dass Eltern mit Zuwanderungsgeschichte höhere Bildungsaspirationen, d.h. Wünsche und Erwartungen hinsichtlich der Bildungsabschlüsse ihrer Kinder, angeben als Eltern ohne Zuwanderungsgeschichte (Dollmann, 2021; Dumont et al., 2013). Außerdem finden sich empirische Hinweise dafür, dass in Familien mit Zuwanderungsgeschichte bei gleichen Leistungen häufiger anspruchsvollere Bildungswege gewählt werden als in Familien ohne Zuwanderungsgeschichte, was auch als „optimistische Bildungsentscheidungen“ bezeichnet wird (Dollmann & Weißmann, 2020). Unklar ist jedoch, inwieweit Bildungsaspirationen und Bildungsentscheidungen zusammenhängen und welche Bedeutung optimistische Bildungsentscheidungen für Bildungsbiografien von Schüler:innen mit Zuwanderungsgeschichte haben. Zahlen sich optimistische Bildungsentscheidungen langfristig aus? Oder stellen sie ein Risiko dar, wenn die Schüler:innen am gewählten Bildungsweg scheitern? Bisher gibt es dazu nur wenige empirische Befunde, diese fanden sowohl Vorteile als auch Risiken (Birkelund, 2020; Dollmann & Weißmann, 2020).

Die vorliegende Studie beschäftigt sich mit Bildungsaspirationen und (optimistischen) Bildungsentscheidungen in Familien mit und ohne Zuwanderungsgeschichte am Beispiel des Schulübergangs in die Sekundarstufe I in Berlin. Dort erhalten Eltern eine Förderempfehlung, sind jedoch nicht an diese gebunden. Optimistische Bildungsentscheidungen lassen sich hier also als Abweichung der tatsächlichen Schulform von der Förderempfehlung operationalisieren, d.h. als Besuch des Gymnasiums ohne entsprechende Empfehlung. Ein Indikator dafür, ob die Entscheidung erfolgreich war, ist das (Nicht-)Bestehen des Probejahrs am Gymnasium. Genauer untersucht werden folgende Fragestellungen:

1. Sagen elterliche Bildungsaspirationen in Familien mit und ohne Zuwanderungsgeschichte optimistische Bildungsentscheidungen vorher?
2. Wählen Eltern mit Zuwanderungsgeschichte häufiger das Gymnasium ohne entsprechende Empfehlung?
3. Bestehen Schüler:innen ohne Gymnasialempfehlung das Probejahr seltener als Schüler:innen mit Empfehlung? Spielt dabei die familiäre Zuwanderungsgeschichte eine Rolle?

Methoden

Die Daten kommen aus dem BERLIN-Projekt (Maaz et al., 2013; Neumann et al., 2017). Die untersuchte Stichprobe bestand aus 1730 Schüler:innen, die zu Beginn der Studie eine 6. Klasse einer Berliner Grundschule besuchten. Etwa 41% von ihnen kamen aus Familien mit Zuwanderungsgeschichte, d.h. ein (n = 316) bzw. beide Elternteile (n = 397) waren nicht in Deutschland geboren. Der Datensatz enthielt Informationen aus einem Elternfragebogen zu elterlichen Bildungsaspirationen (Wahrscheinlichkeit, mit der das Kind das Abitur schafft, 1 Item) sowie zu Familie und Kind (HISEI, elterliche Bildungsstand, Geschlecht Kinde) sowie Informationen

aus Zeugnissen/Schuldokumenten zu Schulnoten, Förderempfehlung, tatsächlich besuchter Schulform und Bestehen des Probejahrs am Gymnasium.

Ergebnisse

Erste Analysen zeigen, dass elterliche Bildungsaspirationen unter Kontrolle von sozioökonomischen und Bildungshintergrund sowie Schulnoten einen signifikanten Prädiktor (*OR* =1.79, $p < .001$) für eine optimistische Bildungsentscheidung darstellen, unabhängig von der Zuwanderungsgeschichte der Familien (Moderationsanalyse, $p > .05$). Außerdem zeigte sich wie erwartet, dass Kinder aus Familien mit Zuwanderungsgeschichte häufiger ohne Förderempfehlung das Gymnasium besuchten, dies galt sowohl für Kinder mit einem (*OR* =2.39, $p = .001$) als auch mit zwei im Ausland geborenen Elternteilen (*OR* =3.53, $p < .001$). Die häufigere Gymnasialwahl ohne Empfehlung in Familien mit Zuwanderungsgeschichte wurde dabei teilweise durch die hohen Bildungsaspirationen vermittelt. Hinsichtlich des Bestehens des Probejahres ergaben sich komplexe Befunde. Von den Schüler:innen aus Familien mit zwei im Ausland geborenen Elternteilen *ohne Gymnasialempfehlung* schaffte knapp die Hälfte (47%) das Probejahr nicht, von denjenigen *mit Gymnasialempfehlung* 17%. Im Vergleich dazu waren es unter den Schüler:innen mit einem im Ausland geborenen Elternteil ohne Gymnasialempfehlung 25%, unter denen mit Empfehlung 9%. Von den Schüler:innen ohne familiäre Zuwanderungsgeschichte konnten 9% bzw. 3% das Probejahr nicht erfolgreich abschließen.

Diskussion

Die Befunde bestätigen unsere Annahmen zum Zusammenhang von Bildungsaspirationen und Bildungsentscheidungen. Zur Frage, ob sich optimistische Bildungsentscheidungen auszahlen, ergab sich ein komplexes Bild. Zwar waren die Erfolgsquoten im Probejahr bei den Schüler:innen mit optimistischen Bildungsentscheidungen geringer – insbesondere bei Schüler:innen mit familiärer Zuwanderungsgeschichte. Andererseits konnte ein erheblicher Anteil der Schüler:innen ohne Gymnasialempfehlung das Probejahr erfolgreich meistern. Die Befunde haben theoretische wie praktische Implikationen, z.B. für Informationsgespräche mit Eltern zur Schulwahl oder möglichen Unterstützungsbedarf bestimmter Schüler:innen für eine erfolgreiche Gymnasiallaufbahn.

Ethnisch diverse Schulklassen begünstigen die soziale Integration geflüchteter Schüler*innen: Ergebnisse einer sozialen Netzwerkanalyse auf Basis repräsentativer Daten

Zsófia Boda[1], Georg Lorenz[2,4], Malte Jansen[2,4], Petra Stanat[2,3,4], Aileen Edele[3,4]
[1]University of Essex, United Kingdom; [2]Institut zur Qualitätsentwicklung im Bildungswesen (IQB); [3]Berliner Institut für empirische Integrations- und Migrationsforschung (BIM); [4]Humboldt-Universität zu Berlin

Aktuell besucht eine beträchtliche Zahl geflüchteter Schüler*innen das deutsche Bildungssystem. Junge Geflüchtete benötigen nicht nur Zugang zu formaler Bildung; auch positive Beziehungen zu Gleichaltrigen sind bedeutsam, da diese eng mit Adaptationsprozessen verknüpft sind (Suárez-Orozco et al. 2018). So begünstigt eine gelungene soziale Integration die psychologische Anpassung (Motti-Stefanidi et al. 2020), Gesundheit (Östberg & Modin 2008) und den Bildungserfolg (Stadtfeld et al. 2019). Welche Faktoren die soziale Integration geflüchteter Schüler*innen unterstützen ist daher eine wichtige Forschungsfrage. Ethnische Diversität in der Schulklasse könnte die Sozialintegration geflüchteter Schüler*innen begünstigen, da Minoritätsschüler*innen grundsätzlich positivere Einstellungen gegenüber Geflüchteten haben dürften (Benner et al., 2018). Gründe hierfür könnten

geringere gruppenbezogene Ängste (Binder et al., 2009) und ähnliche Herausforderungen im Schulkontext sein. Zusätzlich könnte der vermehrte Kontakt mit Minoritätsschüler*innen in ethnisch diversen Klassenzimmern die Vorurteile von Majoritätsschüler*innen gegenüber Geflüchteten reduzieren (Ellison, Shin & Leal 2011), positivere Einstellungen gegenüber Minoritätsschüler*innen begünstigen (Pettigrew & Tropp, 2006) und letztlich ihre Neigung erhöhen, sich mit Geflüchteten anzufreunden.

Die vorliegende Studie untersuchte anhand sozialer Netzwerkanalysen die sozialen Beziehungen geflüchteter Jugendlicher in ihrer Schulklasse. Ziel war es, zu klären, ob a) sich die soziale Integration geflüchteter Schüler*innen von der Integration von Majoritätsschüler*innen bzw. anderen Minoritätsschüler*innen unterscheidet und b) ob und warum Geflüchtete, die ethnisch diverse Schulklassen untersuchen, sozial besser integriert sind als Geflüchtete in weniger diversen Schulkontexten. Da die Verteilung Geflüchteter auf Kommunen und Schulkontexte weitestgehend unabhängig von ihren persönlichen Charakteristika und vielen Kontextmerkmalen, inklusive der ethnischen Komposition der Schulklasse, erfolgte, erlauben unsere Analysen Aussagen darüber, wie der Schulkontext die Integration von neuzugewanderten Geflüchteten beeinflusst.

Als Datenbasis dienten die repräsentativen Daten des IQB-Bildungstrends 2018 (Stanat et al., 2019), in denen 487 Geflüchtete in 304 neunten Klassen identifiziert werden konnten (6.390 Neuntklässler*innen insgesamt in diesen Klassen); insgesamt wurden vollständige Freundschaftsnetzwerke von 39.154 Neuntklässler*innen in 1.807 Schulklassen bestimmt. Als Indikatoren der sozialen Integration wurden Freundschaftsnominierungen sowie die Ablehnung als Sitznachbar*in herangezogen. Als multivariate Analysen, welche darauf abzielen, die Assoziationen zwischen dem Einwanderungsstatus der Schüler*innen und ihren sozialen Beziehungen unabhängig von der Opportunitätsstruktur zu bestimmen, wurden lineare Regressionen für Netzwerkdaten berechnet (Multiple Regression Quadratic Assignment Procedure, MRQAP) (Dekker et al. 2007).

Die Analysen ergeben, dass geflüchtete Jugendliche insgesamt weniger gut sozial integriert sind als Jugendliche, die anderen Minoritätsgruppen bzw. der Majoritätsgruppe angehören. Sie haben weniger Freund*innen und werden häufiger als Sitznachbar*innen abgelehnt. Zudem zeigen die Ergebnisse, dass ethnische Diversität in Schulklassen die Sozialintegration Geflüchteter begünstigt: in ethnisch diversen Klassen haben geflüchtete Schüler*innen tendenziell etwas mehr Freunde, vor allem aber werden sie seltener als Sitznachbar*innen abgelehnt. Während Geflüchtete in ethnisch wenig diversen Klassen im Durchschnitt von 7.48 Mitschüler*innen als Sitznachbar*innen abgelehnt werden, werden sie in Klassen mit mittlerer bzw. hoher Diversität nur von 5.94 bzw. 4.32 Mitschüler*innen abgelehnt. Sowohl die deskriptiven als auch die multivariaten Analysen weisen darauf hin, dass dieser Effekt auf zwei grundlegenden Mechanismen beruht. Erstens haben geflüchtete Lernende in diversen Klassen höhere Chancen, auf Schüler*innen zu treffen, die ebenfalls einer ethnischen Minorität angehören, und diese Schüler*innen akzeptieren geflüchtete Schüler*innen eher als Majoritätsschüler*innen dies tun. Zweitens akzeptieren Majoritätsschüler*innen, die eine ethnisch diverse Klasse besuchen, geflüchtete Schüler*innen eher als Majoritätsschüler*innen in weniger diversen Klassen.

Die Studie liefert wichtige Hinweise darauf, wie unterschiedliche Schulkontexte die Sozialintegration geflüchteter Schüler*innen beeinflussen. Allerdings kann daraus nicht geschlossen werden, dass die Beschulung in diversen Klassen für geflüchtete Schüler*innen insgesamt am günstigsten ist, da für andere wichtige Indikatoren einer gelungenen Integration, etwa die Sprachentwicklung, eine niedrige Diversität günstiger sein könnte.

Mehr als die binäre Optik– Testung eines disparitätsgewichteten Indikators kultureller Diversität

Merle-Sophie Thielmann, Karina Karst, Benedict C. O. F. Fehringer
Universität Mannheim, Deutschland

Theoretischer Hintergrund: Vor dem Hintergrund einer zunehmenden Diversität der Gesellschaft, die auch in Schulen eine neue Normalität darstellt, sind die Effekte kulturell diverser Klassenzusammensetzung, insbesondere auf das Outcome Schülerleistung, in den Fokus der Bildungsforschung gerückt (van Ewijk & Sleegers, 2010; Gomolla, 2021). Von kultureller Diversität werden dabei trotz gemischter Vorhersagen überwiegend negative Effekte auf Schülerleistung erwartet: So stelle eine kulturell diverse Klasse zusätzliche Anforderungen an Lehrkräfte, wenn es darum geht, kulturelle Vorstellungen und Sprachfähigkeiten ihrer SchülerInnen zu berücksichtigen (Braster & Dronkers, 2015). Auch könne das Zugehörigkeitsgefühl von Schüler:innen in Klassen mit weniger Mitschüler:innen ähnlicher kultureller Herkunft niedriger ausfallen, mit negativen Auswirkungen auf Motivation und Leistung (Rjosk et al., 2017). Bei näherer inhaltlicher Auseinandersetzung mit dem Begriff kultureller Diversität wird die Komplexität des Konstruktes deutlich. Je nach theoretischer Betrachtungsweise kann kulturelle Diversität unterschiedliche Aspekte der Klassenzusammensetzung über die Herkunftsländer hinaus fokussieren, wie die Unterschiedlichkeit von Vorstellungen und kulturellen Werten zwischen Ländern. In bisherigen Diversitätsmaßen bleibt diese Komplexität weitgehend unberücksichtigt (Schaeffer, 2016).

Fragestellung: Der Beitrag prüft den Kompositionseffekt kultureller Diversität, operationalisiert durch unterschiedliche Diversitätsmaße, auf die individuelle Leseleistung von SchülerInnen. Auf Basis des bisherigen Forschungsstands und theoretischer Überlegungen erwarten wir einen negativen Effekt. Neben etablierten Maßen—dem Migrationsanteil, der Länderanzahl und den Diversitätsindizes Shannons *H* und Simpsons *D*—messen wir kulturelle Diversität durch ein kulturgewichtetes Maß. Diese in der Soziologie erprobte (Schaeffer, 2013) Erweiterung von Simpsons *D* bezieht neben Anzahl und Verteilung der Herkunftsländer auch deren kulturelle Ähnlichkeit ein (Baldwin & Huber, 2010). Das Maß soll hier erstmals auf schulische Kontexte angewendet werden.

Methode: Als Datengrundlage diente eine Stichprobe von N=1614 Fünftklässler:innen verschiedener Schulformen in Baden-Württemberg. Neben der Leseleistung wurden zu Schuljahresbeginn die Herkunft der Schüler:innen über den Geburtsort beider Eltern erfasst, aus dem analog zum Statistischen Bundesamt (2020) Migrationshintergrund und Herkunft codiert und die Diversitätsmaße berechnet wurden. Zur Bildung des kulturgewichteten Diversitätsmaßes wurden die klassenspezifischen Anteile der Herkunftsländer je paarweise mit einem Disparitätsgewicht multipliziert, das die kulturelle Ähnlichkeit der Länder abbildet, und anschließend summiert (Schaeffer, 2016). Als Gewichte dienten die standardisierten Differenzen der Länder auf den Kulturdimensionen

Secularism und *Self-Expression* nach Inglehart und Welzel (2005), welche auf Basis einer Subskala zur Wichtigkeit der Vermittlung kultureller Werte an Kinder aus der World Values Survey (Haerpfer et al., 2022) ermittelt wurden (Schaeffer, 2013). Die Diversitätsmaße gingen als Prädiktoren in getrennte Mehrebenenmodelle ein. Als abhängige Variable diente jeweils die Leseleistung zu Ende des ersten Schulhalbjahres, erfasst durch eine adaptierte Version des *Lernstand 5*. Kontrolliert wurde für die Eingangsleistung, kulturelles Kapital, Geschlecht, durchschnittliche Klasseneingangsleistung und den kulturellen Hintergrund der Schüler:innen. Zur Berechnung wurde nach Imputation fehlender Werte im Paket mice das Paket lme4 genutzt.

Ergebnisse: Es zeigte sich ein signifikant negativer Haupteffekt etablierter kultureller Diversitätsmaße auf die Leseleistung ($-0.11 \leq \beta \leq -0.06$). Die etablierten Maße untereinander zeigten hohe, signifikante Korrelationen ($0.47 \leq r \leq 0.85$). Das kulturgewichtete Diversitätsmaß korrelierte dagegen, auch im Vergleich zu anderen Studien (Schaeffer, 2013), relativ gering mit anderen Diversitätsmaßen, insbesondere mit der Länderanzahl ($r = 0.23$, $p < .001$) und Shannons H ($r = 0.47$, $p < .001$). Klassen mit ähnlichem Migrationsanteil und sogar mit ähnlichem Diversitätsindex unterschieden sich also, sobald die Ähnlichkeit der kulturellen Hintergründe als Aspekt der Diversität herangezogen wurde. Auch die kulturgewichtete Diversität war ein signifikanter Prädiktor für die Leseleistung von Schüler:innen ($\beta = -0.07$, $p < .05$). Unsere Ergebnisse weisen darauf hin, dass kulturgewichtete Indizes geeignete Prädiktoren von Leistung, jedoch nicht gleichzusetzen mit anderen gängigen Maßen kultureller Diversität sind. Vielmehr bieten kulturgewichtete Indizes eine Möglichkeit, spezifischer jeweils die Aspekte kultureller Diversität zu operationalisieren, die von theoretischem Interesse sind. So bieten Sie das Potential, differenzierter die Effekte und Wirkung kultureller Diversität im Klassenzimmer zu erklären.

F-02: Digitales Lernen III

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S06 S00 B32
Chair der Sitzung: Michael Sailer

Eine systematische Überblicksarbeit zu Meta-Analysen über digitales Lernen aus einer Lernaktivitätenperspektive

Michael Sailer[1], Rebecca Maier[1], Sonja Berger[1], Tamara Kastorff[1], Karsten Stegmann[2]
[1]LMU München, Germany; [2]Universität Passau, Germany

Digitale Technologien haben das Lehren und Lernen im Allgemeinen und die Hochschullehre im Besonderen in vielerlei Hinsicht verändert. Digitale Technologien haben das Potenzial positive Effekte auf den Lernerfolg von Studierenden zu erzielen (Tamim et al., 2011). Jedoch deuten aktuelle Befunde darauf hin, dass es auf die Art und Weise des digitalen Technologieeinsatzes ankommt und weniger auf die Häufigkeit deren Einsatzes (Fütterer et al., 2022). Meta-Analysen können eine adäquate Antwort geben, wie digitale Technologien eingesetzt werden können, um den Lernerfolg von Studierenden zu fördern. Sie konzentrieren sich jedoch häufig auf die Auswirkungen bestimmter digitaler Technologien und ihrer technologiespezifischen Merkmale auf die Lernergebnisse (Stegmann, 2020). Sie stellen daher selten generelle Erkenntnisse zu den Auswirkungen digitaler Technologien bereit, die sich über verschiedene Technologien hinweg generalisieren lassen. Auf Grund der schnell voranschreitenden Weiterentwicklung digitaler Technologien haben technologiespezifische Erkenntnisse jedoch eine eher geringe Halbwertszeit. Für die Hochschulpraxis werden stattdessen evidenzorientierte Empfehlungen benötigt, die über Spezifika einzelner digitaler Technologien hinausgehen. Das ICAP Modell (Chi & Wylie, 2014), das interaktive, konstruktive, aktive und passive Lernaktivitäten unterscheidet, ermöglicht es unabhängig von der verwendeten Technologie die Art und Weise des digitalen Technologieeinsatzes genauer zu spezifizieren. Auf Basis des ICAP-Modells lässt sich die Hypothese aufstellen, dass die Effekte des technologiegestützten Lernens durch die Lernaktivitäten vermittelt werden, die durch die Technologie ermöglicht werden.

In der vorliegenden systematischen Übersichtsarbeit von Meta-Analysen wurden die Auswirkungen von Lehre mit und ohne digitaler Technologie in der Hochschulbildung untersucht, wobei die Lernaktivitäten der Studierenden unter den Bedingungen mit und ohne Technologie berücksichtigt wurden. Die Zuordnungen von Lernaktivitäten basierten auf dem ICAP-Modell, das postuliert, dass die direkt beobachtbaren Lernaktivitäten eine hinreichende Annäherung an die zugrundeliegenden kognitive Prozesse sind. Bei unserer systematischen Suche nach Meta-Analysen, haben wir N = 29 Effekte aus N = 17 den Kriterien entsprechenden Meta-Analysen identifiziert und in diese Überblicksarbeit einbezogen. Die Meta-Analysen wurden zwischen 2012 und 2022 veröffentlicht.

Die Ergebnisse zeigen, dass sich die Lernleistung der Studierenden nicht signifikant verändert, wenn die Lehre mit digitaler Technologie als Substitution für die Lehre ohne Technologie eingesetzt wurde (d. h. wenn sich die zentralen Lernaktivitäten der Studierenden nicht verändert haben und die Lernaktivitäten nicht zusätzlich unterstützt wurden). Wenn die Lernaktivität gleich blieb, die digitale Technologie jedoch eine lernaktivitätsspezifische Unterstützung kognitiver Lernprozesse bot, zeigen sich positive Effekte des Technologieeinsatzes auf den Lernerfolg. Darüber hinaus hatten digitale Technologien, die zur Steigerung der Lernaktivitäten im Sinne des ICAP Modells eingesetzt wurden, einen positiven Einfluss auf die Lernergebnisse der Studierenden.

Die Ergebnisse stützen die Hypothese, dass die Effekte von technologiegestützter Lehre durch die Lernaktivitäten vermittelt werden. Digitale Technologien wirken sich demnach dann positiv auf das Lernen der Studierenden an der Hochschule aus, wenn sie lernaktivitätsspezifisch kognitive Lernprozesse unterstützen (z. B. durch Feedback oder Scaffolding) oder wenn sie bei den Lernenden zu höheren Lernaktivitäten im Sinne des ICAP Modells führen. Darüber hinaus unterstreichen die Ergebnisse die entscheidende Rolle der Untersuchung von Lernaktivitäten zusätzlich zu spezifischen technologischen Merkmalen im Rahmen von Meta-Analysen. Die bedeutsamen positiven Effekte digitaler Technologien sollten für die Hochschulpraxis zur Konsequenz haben, dass der lernwirksame digitale Technologieneinsatz ein integraler Bestandteil der Lehrveranstaltungen an der Hochschule ist. Die Auswertung der Meta-Analysen hat gezeigt, dass die Implementierung einer breiten Palette von fortgeschrittenen Lernaktivitäten mit Hilfe digitaler Technologien ein effektiver Weg ist, um das Lernen der Studierenden zu verbessern. Voraussetzungen dafür ist jedoch, neben der Bereitstellung der technischen Infrastruktur, insbesondere die hochschuldidaktische Aus- und Fortbildung der Lehrenden.

Einsatz von Videoannotationen zur Förderung der professionellen Wahrnehmung von technologie-gestützten Unterrichtssituationen

Franziska Tschönhens[1], Ulrike Franke[1], Tim Fütterer[1,2], Andreas Lachner[3,4]
[1]Tübingen School of Education, Universität Tübingen; [2]Hector-Institut für Empirische Bildungsforschung und Pädagogische Psychologie, Universität Tübingen; [3]Tübingen Center for Digital Education, Universität Tübingen; [4]Institut für Erziehungswissenschaft, Universität Tübingen

Theoretischer Hintergrund. Der Einsatz von Bildungstechnologien im Unterricht ist mit hohen Anforderungen an Lehrkräfte verbunden, da gleichzeitig das Unterrichtsgeschehen und die Nutzung der Bildungstechnologien beobachtet und reflektiert werden muss. Um kritische Situationen erkennen und bewerten zu können, benötigen (angehende) Lehrkräfte Kompetenzen im Bereich der professionellen Wahrnehmung, die bereits in der universitären Lehrerbildung angebahnt werden sollte (Seidel & Stürmer, 2014). Annotationstools, die die positionsgenaue Markierung und Bewertung von videobasierten, technologie-gestützten Unterrichtssituationen erlauben, besitzen in diesem Kontext besonderes Potential. Um das Potential solcher Tools vollumfänglich auszuschöpfen, benötigt es jedoch zusätzliche Unterstützung (Seidel, 2013). In der instruktionspsychologischen Forschung werden daher Prompts (Nückles, 2009) oder auch Modellierungsbeispiele (Renkl, 2014) diskutiert, um Lernende bei der Bearbeitung von multimedialen Inhalten in ihrem Lernprozess zu unterstützen. Allerdings wurden Videoannotationen im Vergleich zu komplexen Unterrichtssituationen bisher ausschließlich in salienten Kontexten (z.B. individuelle Bewegungsanalysen) untersucht. Deshalb ist bislang offen, wie und ob diese Unterstützungsstrategien in Annotationsumgebungen gewinnbringend eingesetzt werden können, dass sie die professionelle Wahrnehmung in technologie-gestützten Unterrichtssettings fördern.

Fragestellungen. In der vorliegenden präregistrierten experimentellen Feldstudie werden zwei Instruktionsstrategien (Modellierungsbeispiele und Prompts) bei Videoannotationen systematisch eingesetzt, um die professionelle Wahrnehmung von technologie-gestütztem Unterricht bei Lehramtsstudierenden zu fördern. Dabei werden folgende Fragestellungen untersucht:

(F1) Welchen Effekt haben Prompts und Modellierungsbeispiele auf das Erkennen problematischer Unterrichtssituationen (Noticing)?

(F2) Welchen Effekt haben Prompts und Modellierungsbeispiele auf die Bewertung von problematischen Unterrichtssituationen (Reasoning)?

Zusätzlich untersuchten wir explorativ Interaktionseffekte von Prompts und Modellierungsbeispielen auf das Noticing und Reasoning von Lehramtsstudierenden?

Methode. Die feld-experimentelle Studie wurde mit einem 2x2x3-faktoriellen Design mit dem Vorhandensein von Modellierungsbeispielen (ja, nein) und Prompts (ja nein) als Zwischensubjektfaktoren, und Zeit (Video1-3) als Innersubjektfaktor im Kontext einer Master-Lehramtsvorlesung zum Lehren und Lernen mit digitalen Medien umgesetzt (N = 167, n = 122 vollständige Datensätze). Alle Lehramtsstudierende füllten zunächst einen kurzen Fragebogen zu demografischen Daten und deren technologisch-pädagogisches Vorwissen (Lachner et al., 2019) aus. Danach schauten sie in der Annotationssoftware *LiveFeedback plus* drei Videos von Unterrichtssequenzen, deren Aufbau identisch war, an. Dabei wurden sie gebeten, kritische Situationen zu markieren und zu kommentieren. Das erste Video diente als Prätest der professionellen Wahrnehmung. Zwischen der ersten und zweiten Video erhielten die Studierenden randomisiert entweder ein videobasiertes Modellierungsbeispiel, welches adäquate Kommentierungsstrategien modellierte oder nicht. Zudem erhielten diese randomisiert während des zweiten Videos Prompts zur Aktivierung von adäquaten Kommentierungsstrategien oder nicht. Die Annotationen wurden hinsichtlich der erkannten kritischen Unterrichtssituationen (Noticing) und der adäquaten Bewertung der erkannten Unterrichtssituation (Reasoning) ausgewertet (ICC = .978; Martin et al., 2022; Seidel & Stürmer, 2014). Um die Fragestellungen zu überprüfen, wurden Kovarianzanalysen (ANCOVAs) mit Messwiederholungen (Video 1-3) unter Berücksichtigung des Vorwissens als Kovariate durchgeführt.

Ergebnisse und ihre Bedeutung. Hinsichtlich Noticing zeigte sich ein signifikanter Abfall über die Zeit, $F(2,234) = 3.353$; $p = .037$; $\eta^2 = .028$. Keiner der anderen Effekte war signifikant. Hinsichtlich Reasoning zeigte sich ebenfalls ein signifikanter Abfall über die Zeit $F(2,234) = 3.958$; $p = .020$; $\eta^2 = .033$. Der Interaktionseffekt zwischen Modellierungsbeispielen und Zeit war signifikant $F(2,234) = 13.381$; $p < .001$; $\eta^2 = .103$, da Studierende mit Modellierungsbeispiel die Unterrichtssituationen im zweiten Video besser bewerten konnten als ohne Beispiel. Keiner der anderen Effekte war signifikant. Die vorläufigen Befunde zeigen, dass insbesondere Modellierungsbeispiele eine effektive Möglichkeit bieten, die Bewertung von technologie-gestützten Unterrichtssituationen zu verbessern. Prompts zeigten dagegen keinen Effekt. In Folgestudien sollte daher untersucht werden, wie Prompts als minimal-invasive Methode Lehramtsstudierende unterstützen kann sowie Transfereffekte auf andere Unterrichtssituationen verstärkt werden können, um die professionelle Wahrnehmung in technologie-gestützten Unterrichtssituationen weiterzuentwickeln.

Fördert „Problem-solving-first“ den Erwerb von Transferwissen beim Lernen mit Erklärvideos?

Maria Wevers, Martin Hänze
Universität Kassel, Germany

In einer Welt, die von kontinuierlichen teilweise rasanten Veränderungen geprägt ist, ist es von besonderer Wichtigkeit Lernenden Wissen zu vermitteln, welches sie auch in neuen noch unbekannten Situationen nutzen können. Ziel der Bildung sollte demnach nicht die Vermittlung starren Faktenwissen, sondern flexiblen Transferwissen sein (Barnett & Ceci, 2002; Mayer, 2002), das Lernende im Alltag praktisch nutzen können und ihnen ermöglich mit stetigen technischen und wissenschaftlichen Fortschritt umzugehen. Wenngleich die Bedeutsamkeit auf der Hand liegt, ist die Vermittlung von Transferwissen noch immer eine große Herausforderung (Anderson & Beavis, 2018). Der instruktionale Ansatz „Problem-solving-first“ (PF) hat in den vergangenen Jahren jedoch seinen Erfolg hinsichtlich der Förderung von Transferwissen im primär naturwissenschaftlichen Bereich unter Beweis stellen können (Chen & Kalyuga, 2020; Loibl et al., 2017). Bei diesem instruktionalen Ansatz versuchen Lernende zunächst ein komplexes Problem oder eine Anwendungsaufgabe zu lösen. Die Vermittlung der für das Lösen relevanten Informationen und Konzepte erfolgt im Anschluss. Es wird angenommen, dass dieser Ansatz die Lernmotivation sowie kognitive Aktivierung von Lernenden positiv beeinflusst. Darüber hinaus werden im vorangestellten Problemlöse-Prozess bestehende Wissensnetze (Vorwissen) in besonderem Maße aktiviert. Durch diese Mechanismen kann das neue anschließend vermittelte Wissen tiefer verarbeitet und zugrundeliegende Konzepte erkannt werden, sodass Lernende im Anschluss meist besser in der Lage sind auch neuartige Anwendungsaufgaben erfolgreich zu lösen (Loibl et al., 2017; Schwartz & Martin, 2004). Wenngleich das Design bereits bei der Vermittlung naturwissenschaftlicher Inhalte in bestimmten Lernsituationen positive Effekte zeigt (Loibl et al., 2017), ist noch unzureichend geklärt, auf welche weiteren Lernsituationen und welche Inhalte sich die Erfolge übertragen lassen.

In einer ersten Studie wurde daher untersucht, inwieweit PF auch den Transferwissenslernerfolg beim Lernen mit Erklärvideos steigern kann. Denn Studien verweisen darauf, dass beim Lernen mit Erklärvideos eine unzureichende Verarbeitungstiefe erreicht wird und demnach auch Lernerfolge in Bezug auf Transferwissen ausbleiben können. Erklärt wird dies damit, dass Erklärvideos bei Lernenden zu geringerer kognitiver Aktivierung (Paik & Schraw, 2013) sowie eine hohe kognitive Belastung (Leahy & Sweller, 2016) führen können. PF könnte genau jene Schwächen, sodass auch im Zusammenhang mit Erklärvideos PF zur Förderung von Transferwissen führt. Zur Überprüfung der Annahme haben 58 Studierende im Rahmen einer Physik-Grundschullehramtsvorlesung an der quasi-experimentellen Studie teilgenommen. Die Studierenden rezipierten ein Erklärvideo zum Thema Licht und Schatten entweder vor oder nach dem Bearbeiten einer komplexen Anwendungsaufgaben. Die Ergebnisse zeigen, dass Lehramtsstudierende höhere Lernerfolge hinsichtlich des Transferwissens erzielen, wenn sie gemäß des PF-Ansatzes kooperativ mit dem Erklärvideo lernen.

In einer zweiten Studie sollte untersucht werden, inwieweit der PF-Ansatz auch positive Effekte auf das Lernen von pädagogisch-psychologischen Inhalten entfalten kann. Da in diesem Themenfeld Problemlösungen und Aufgaben anders gestaltet sind und es selten eine eindeutige richtige Lösung gibt, ist nicht sicher, ob die genannten PF-Wirkmechanismen auch hier gelten. Zur Prüfung wurde ebenfalls eine quasi-experimentelle Studie durchgeführt, bei der Lehramtsstudierende an zwei Terminen, entweder vor oder nach dem Rezipieren pädagogisch-psychologischer Fachinhalte eine komplexe Anwendungsaufgabe bearbeiteten. Es nahmen 209 Studierende in der ersten Woche zum

Thema „Cognitv Load Theorie“ teil und 199 Studierende in der zweiten Woche zum Thema „Erwartungs-mal-Wert Modell“. Die Ergebnisse dieser Untersuchung befinden sich derzeit noch in der Auswertung, werden jedoch zum Zeitpunkt der GEBF vorliegen.

Die Erkenntnisse dieser beiden Studien ermöglichen es, praktisch hochrelevante Empfehlungen abzuleiten, wie insbesondere der Erwerb von Transferwissen beim Lernen mit Erklärvideos nicht nur in naturwissenschaftlichen Themenfeldern gefördert werden kann.

F-03: Unterrichtsqualität, Instruktionsdesign und epistemische Emotionen

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S06 S00 B41
Chair der Sitzung: Robin Göller

Zur Bedeutung epistemischer Emotionen für den Aufbau fachlichen und fachdidaktischen Wissens von Lehramtsstudierenden im Fach Mathematik

Robin Göller[1], Lars Jenßen[2], Katja Eilerts[2], Michael Besser[1]
[1]Leuphana Universität Lüneburg, Germany; [2]Humboldt-Universität zu Berlin

Theoretischer Hintergrund. Fachliches und fachdidaktisches Wissen werden zusammen mit pädagogisch-psychologischem Wissen als zentrale Aspekte des Professionswissens von Lehrkräften angesehen (Baumert & Kunter, 2006; Shulman, 1987; Voss et al., 2011). Derartiges Professionswissen ist zentraler Bestandteil der professionellen Kompetenz von Lehrkräften und hat einen empirisch nachgewiesenen Einfluss auf Unterrichtsqualität und Leistungen von Schüler:innen (Baumert et al., 2010; Hill et al., 2005; Krauss et al., 2020). Darüberhinausgehend wird zunehmend auch die Bedeutung von (epistemischen) Emotionen für den Wissenserwerb von Schüler:innen und (zukünftigen) Lehrkräften diskutiert (Pekrun, 2021). Epistemische Emotionen sind Emotionen, die beim Abgleich zwischen neuen Informationen und bestehenden Überzeugungen und Wissensstrukturen entstehen (Muis et al., 2018) und korrelieren mit Lernstrategien, Werten und Leistungen von Lernenden (Camacho-Morles et al., 2021; Muis et al., 2015; Pekrun et al., 2017). Der vorliegende Beitrag fokussiert auf die epistemischen Emotionen Neugierde und Frustration und untersucht, welchen Effekt diese auf den Erwerb fachlichen und fachdidaktischen Wissen von Lehramtsstudierenden im Fach Mathematik haben. Konkret wird folgende Fragestellungen untersucht:

Fragestellung. Welchen Einfluss haben die epistemischen Emotionen Neugierde und Frustration auf das mathematische und mathematikdidaktische Wissen von Lehramtsstudierenden?

Methode. Analysiert wurden Daten von 145 Mathematik-Lehramtsstudierenden (112 weiblich, 33 männlich), die im Sommer 2022 an einer Paper-Pencil-Erhebung teilgenommen haben. Neben der Erfassung mathematischen Fachwissens sowie mathematikdidaktischen Wissens mittels standardisierter Testinstrumente (siehe u. a. Besser et al., 2021) wurden ebenso mathematikbezogene Neugierde (ω= .77) und Frustration (ω=.80) über jeweils drei Items mithilfe fünfstufiger Likert-Skalen erhoben (Pekrun et al., 2017). Darüber hinaus wurden als Kontrollvariablen u. a. die Semesteranzahl, das Geschlecht, die Abiturnote sowie die letzte Schulnote in den Fächern Mathematik und Deutsch (1=sehr gut, 6=ungenügend) erfragt. Zur Datenanalyse wurde ein Strukturgleichungsmodell mit den Testscores der beiden Wissenstests als endogenen Variablen und den anderen genannten Konstrukten als exogenen Variablen mithilfe des R-Packages lavaan berechnet.

Ergebnisse. Das betrachtete Strukturgleichungsmodell zeigt insgesamt gute Modell-Fit-Werte (χ^2=50.69; *df*=46; *p*=.29; *CFI*=0.987, *TLI*=0.981, *RMSEA*=0.027) und erklärt 41% der Varianz des mathematischen Fachwissens sowie 32% der Varianz des mathematikdidaktischen Wissens. Neugierde hat einen signifikant positiven Effekt auf das mathematische Wissen (standardisiertes *β*=0.207, *p*=.010), Frustration einen signifikant negativen Effekt (*β*=-0.162, *p*=.042). Auch für das mathematikdidaktische Wissen ist der Effekt von Neugierde signifikant positiv (*β*=0.228, *p*=.007), der

Effekt von Frustration ist allerding nicht signifikant (β=-0.150, p=.077). Signifikante Effekte der miterfassten Kontrollvariablen auf das mathematische Fachwissen finden sich für die Abiturnote (β=-0.327, p<.001), die letzte Schulmathematiknote (β=-0.166, p=.029) sowie das Geschlecht (β=0.408, p<.001) zugunsten der männlichen Teilnehmer. Auf das mathematikdidaktische Wissen haben von den Kontrollvariablen die Semesteranzahl (β=0.338, p<.001) sowie die Abiturnote (β=-0.249, p=.001) einen signifikanten Effekt.

Diskussion. Die Ergebnisse zeigen signifikante Effekte beider betrachteten Emotionen auf das mathematische Fachwissen. Neugier hat auch für das mathematikdidaktische Wissen einen positiven Effekt. Darüberhinausgehend ist mit Blick auf die Ausbildung von Professionswissen in der Lehramtsausbildung kritisch zu diskutieren, dass Vorwissen und Geschlecht für das mathematische Fachwissen offenbar eine größere Bedeutung haben als die Semesteranzahl. Auf das mathematikdidaktische Wissen (das zu Studienbeginn ggf. eher weniger ausgebildet ist) haben eine höhere Semesteranzahl sowie eine bessere Abiturnote einen positiven Effekt (unter Kontrolle der anderen betrachteten Variablen). Bei der Interpretation der Ergebnisse muss die relativ kleine Stichprobe berücksichtigt werden (weitere standortübergreifende Erhebungen sind geplant). Zudem können die Operationalisierungen der Testinstrumente und die Testsituation (z. B. aufgrund der Anonymisierung kein extrinsischer Anreiz gut abzuschneiden) die Ergebnisse möglicherweise beeinflusst haben.
Insgesamt unterstreichen die Ergebnisse jedoch die Bedeutung von Emotionen für das Professionswissen von Lehramtsstudierenden, die in der Diskussion zur Ausbildung professioneller Kompetenzen von Lehrkräften ggf. noch stärker in den Blick genommen werden sollte (für eine tiefere Diskussion siehe Pekrun, 2021).

Unterrichtsplanung mit virtuellen Lernsimulationen – Ein sinnvoller Ansatz zur Förderung des Einsatzes automatisierter Lernprompts durch angehende Lehrkräfte?

Ascha Ahmed, Sabrina Ludwig, Manuel Vogler, Lucas Lörch, Carmela Aprea, Viola Deutscher
Universität Mannheim, Deutschland

Theoretischer Hintergrund

Die Unterrichtsplanung nimmt neben der Unterrichtsdurchführung und -kontrolle eine essentielle Rolle für Lehrkräfte ein (Peterßen, 2000) und kann als Designaufgabe über folgende Schritte konzipiert werden (Aprea et al., 2021):

- Domänenspezifische und kognitive Aufgabenanalyse
- Bestimmung der Lernziele und Lerninhalte
- Konfiguration methodischer Gestaltungsparameter

In Bezug auf den dritten Schritt lassen sich im Rahmen von digitalen Lernumgebungen zunehmend auch KI-Elemente und hier insbesondere automatisierte Lernprompts als Gestaltungselemente implementieren. Lernprompts sind Hilfestellungen bzw. Hinweise, in Form von Fragen, Vorschlägen und Feedback, die während des Lernprozesses erscheinen und die Anwendung relevanter Verarbeitungsstrategien fördern (Wirth, 2009). Diese können als personalisiertes Gestaltungselement zur Implementierung individueller Förderung in den berufsschulischen Unterricht genutzt werden

(Deutscher et al., 2022). Didaktisch kann dabei an den 4C/ID-Ansatz (van Merrienböer, 2017) angeknüpft werden, indem Lernprompts als individuelles scaffolding and fading konzipiert werden.

Mithilfe der Integration der virtuellen Bürosimulation LUCA (Rausch et al., 2021) in ein bestehendes Unterrichtsplanungs-Modul an der Universität Mannheim wurden Masterstudierende im Fach Wirtschaftspädagogik im Frühjahrs-/Sommersemester 2022 im Rahmen der Projektinitiative „Innovation in der Hochschullehre“ auf die Herausforderungen der Digitalisierung im Lehrberuf vorbereitet. Ziel des Projekts war es zum einen, die Studierenden dazu zu befähigen, Virtual Learning Simulations (VLS) zu erstellen, die einen hohen Immersionsgrad aufweisen (Braunstein et al., 2021). Zum anderen sollte mit Hilfe dieses Virtual-Learning-Formats die Fähigkeit der Studierenden gefördert werden, Logdaten-basierte Lernprompts sinnvoll für ihren Unterricht einzusetzen.

Fragestellung

Mit dem Einsatz der virtuellen Bürosimulation im Modul ergaben sich folgende Forschungsfragen:

1. Unterscheidet sich das Wissen zum Thema Lernprompts sowie die selbstberichtete Selbstsicherheit im Einsatz der Lernprompts vor und nach der Teilnahme an der Seminarveranstaltung?
1. Verändern sich Einstellungen und Absichten in Bezug auf den Einsatz von Lernprompts für den Unterricht?

Methode

Mithilfe eines Prä- und Posttests mit Skalenfragen mit 5-stufiger Likerts-Skala, Multiple-Choice-Items sowie offenen Testfragen wurde das Wissen der Studierenden zu Beginn (n_Prätest = 35) und am Ende (n_Posttest = 28) der Lehrveranstaltung zu Lernprompts sowie eine Einschätzung des potentiellen Einsatzes dieser im Unterricht erfasst.

Ergebnisse

Die Studierenden schätzten nach dem Seminar ihr Wissen zum Thema Lernprompts signifikant höher ein als vor dem Seminar (“Ich habe Kenntnisse über die Einsatzmöglichkeiten von Lernprompts im Unterricht oder in der betrieblichen Ausbildung.” $F(1, 63) = 106{,}80$; $p < .001$; partial $\eta^2 = 0{,}63$). Gleiches galt für die Selbstsicherheit im Einsatz von Lernprompts (“Im Einsatz von Lernprompts im Unterricht oder in der betrieblichen Ausbildung fühle ich mich selbstsicher.” $F(1, 63) = 65{,}95$; $p < .001$; partial $\eta^2 = 0{,}51$). Die Intention, Lernprompts aktuell im Unterricht oder der betrieblichen Ausbildung einzusetzen, stieg leicht (“Ich plane aktuell Lernprompts in meinem Unterricht oder in meiner betrieblichen Ausbildung einzusetzen.” $F(1, 63) = 9{,}52$; $p < .01$; partial $\eta^2 = 0{,}13$). Gleichbleibend hoch war die Zustimmung zur Aussage, dass der Einsatz von Lernprompts das Potential hat, den Unterricht oder die betriebliche Ausbildung zu verbessern (“Der Einsatz von Lernprompts hat das Potential meinen Unterricht oder betriebliche Ausbildung zu verbessern.” $F(1, 63) = 2{,}26$; $p = 0.14$). Relativ nahm die Anzahl der Studierenden, die der Nutzung von Lernprompts aktuell zustimmen würden über beide Messzeitpunkte zu (Prätest = 65,71 %; Posttest = 85,71 %). Auch der objektive Wissenstest in Form von MC- und offenen Antworten deutet auf einen signifikanten Wissenserwerb hin (z.B. “Was sind Lernprompts?” max. 2 Punkte. $F(1, 63) = 7{,}55$; $p < .01$; partial $\eta^2 = 0{,}11$).

Die Ergebnisse zeigen, dass das Wissen zum Thema Lernprompts und die Intention zu deren Einsatz im Laufe des Seminars signifikant zunahm. Dies ist ein erster Hinweis darauf, dass die Vermittlung von Unterrichtsplanung mit virtuellen Lernumgebungen ein geeignetes didaktisches Mittel darstellen kann, um die Arbeit mit personalisierten Lernprompts zu fördern - wobei ein Kontrollgruppendesign aktuell noch aussteht.

Erfassung des Potentials zur kognitiven Aktivierung in Aufgaben

Anna Heinle, Anja Schiepe-Tiska, Doris Lewalter
Technische Universität München, Deutschland

1.Theorie

Kognitive Aktivierung (KA) gilt als zentrales Unterrichtsqualitätsmerkmal mit dem Potenzial, Ergebnisvariablen wie Schülerleistungen oder Motivation zu erklären (u.a. Baumert et al., 2010; Fauth et al., 2014). KA ist theoretisch in kognitiv-konstruktivistische Lerntheorien (u.a. Vygotsky, 1978) sowie Conceptual-Change-Theorien (u.a. Schneider & Hardy, 2013) verankert und wird als Anregung eines tieferen und konzeptuellen Verständnisses von Inhalten definiert (Klieme et al., 2009). Eine konsistente Operationalisierung von KA stellt bisher jedoch eine zentrale Herausforderung dar (s. Prateorius et al., 2018). Aktuelle Befunde (u.a. Drollinger-Vetter, 2011; Praetorius et al., 2014), aber auch theoretische Diskussionen plädieren für eine fachliche Spezifizierung dieses Qualitätsmerkmals (u.a. Praetorius & Gräsel, 2021). Empirische Untersuchungen zur generischen und fachspezifischen Dimensionalität des Konstrukts fehlen bisher.
Bislang wird KA verstärkt anhand Unterrichtsvideos erfasst. Eine insgesamt wenig genutzte Alternative stellt die Analyse von Aufgaben dar, obwohl diese besonders im Mathematik- und Naturwissenschaftsunterricht eine wichtige Funktion in der Initiierung von Lernprozessen einnehmen (u.a. Knoll, 2003).

Ziel des Beitrags ist es, anknüpfend an die aktuelle Diskussion (Praetorius et al., 2018; Herbert & Schweig, 2021) ein Kategorienschema zur generischen und fachspezifischen Erfassung des Potenzials zur kognitiven Aktivierung (PKA) in Mathematik- und Naturwissenschaftsaufgaben vorzustellen und empirisch zu überprüfen. Dieses umfasst die theoretisch generische wie fachspezifische Kategorie 1) *Herausfordernde Aufgaben und Fragen* mit den Subkategorien *Lebensweltbezug*, *Offenheit*, *sprachlogische Komplexität*, *mathematisches Modellieren* (Mathematik) und *Experimentieren* (Naturwissenschaften) sowie die generischen Kategorien 2) *Vorwissen aktivieren*, 3) *Denkweisen der Schüler:innen ergründen*, 4) *rezeptives Verständnis von Lernen (neg.)*, 5) d*iskursives/ ko-konstruierendes Lernen* und 6) *Metakognition*.
Folgende Forschungsfragen sollen beantwortet werden:

2. Fragestellungen

1. Inwieweit beinhalten aktuelle Mathematik-, Physik- sowie Chemieaufgaben in deutschen Lehrwerken der 9. Jahrgangsstufe Merkmale von PKA?
2. Inwieweit unterscheiden sich die Merkmale von PKA in den Fachbereichen?
3. Inwieweit lassen sich die latenten Konstrukte *generisches PKA* und *fachspezifisches PKA* empirisch abbilden?
4. Lässt sich das latente Konstrukt *generisches PKA* besser durch ein ein- oder sechsdimensionales Modell abbilden?
5. Lässt sich das theoretisch generische wie fachspezifische latente Konstrukt *Herausfordernde Aufgaben und Fragen* besser durch ein ein- oder zweidimensionales Modell abbilden?

3. Methode

Um ein bestmögliches Abbild aktueller Lehrwerksaufgaben der 9. Jahrgangsstufe deutschlandweit zu erhalten, wurden Lehrwerke der Verlagsgruppen Cornelsen, Klett und Westermann ausgewählt. Beim Sampling wurde mehrstufig sowie nach einem top-down Verfahren vorgegangen und für die Aufgabenauswahl ein Rotationsdesign angewendet. Daraus ergaben sich *N*= 377 Analyseeinheiten (n= 136 Mathematik, n=115 Physik und n=126 Chemie).

Das Kategorienschema wurde deduktiv und zusätzlich induktiv entwickelt (Früh, 2004) und umfasst dichotome Antwortkategorien. Für die Reliabilitätsprüfung wurden die Analyseeinheiten doppelt kodiert. Die Cohens Kappa(κ) Werte der Kategorien liegen zwischen 0.6 und 1 und verweisen damit auf eine starke bis fast vollständige Übereinstimmung (Wirtz & Caspar, 2002). Die empirische Überprüfung der Konstrukte wurde anhand eindimensionaler Raschmodelle realisiert und die Unidimensionalität durch bedingte stochastische Unabhängigkeit überprüft. Für die Modellschätzungen wurde das R-package *pairwise* (Heine, 2021) genutzt.

4. Ergebnisse und ihre Bedeutung

Am häufigsten wurde *Lebensweltbezug* in etwas weniger als 50% aller Aufgaben identifiziert. Die Merkmale *diskursives/ko-konstruierendes Lernen* sowie *Metakognition* wurden kein bzw. maximal zwei Mal kodiert. Zusammenfassend treten die Merkmale von PKA mit nur geringen Fachunterschieden eher selten auf. Somit bleibt das PKA in aktuellen MINT-Schulbuchaufgaben bisher ungenutzt. In allen Fächern konnte die Eindimensionalität der latenten Konstrukte *generisches* und *fachspezifisches PKA* bestätigt werden. Der Vergleich von Bi-Faktor-Modellen und explorativen Faktorenanalysen des Konstrukts *generisches PKA zeigt*, dass es sich besser durch ein eindimensionales Modell abbilden lässt (1dim Modell: 2634.178, 6dim Modell: 2781.695). Das latente Konstrukt *Herausfordernde Aufgaben und Fragen* lässt sich ebenfalls besser durch ein eindimensionales Modell abbilden und zeigt damit, dass die generische und fachspezifische Dimension des PKA in Aufgaben empirisch schwer trennbar ist. Die Autorengruppe plädiert folglich für eine sowohl fächerübergreifende als auch fachspezifisch Intensivierung der Theorieentwicklung des Konstrukts.

Unterrichtsqualität und digitale Medien: Ein experimenteller Vergleich von Texten, Tablets und immersiver VR

Alexander Georg Büssing
Universität Trier, Germany

Innovative Technologien wie die immersive virtuelle Realität (VR) werden seit Jahren untersucht, wobei unterschiedliche Evidenzen für förderliche oder hinderliche Lerneffekte bestehen (Mayer & Makransky, 2022). Dabei zeichnet sich die VR insbesondere durch die erhöhte Immersion aus, die technische Gegebenheiten beschreibt, um realistische Darstellungen zu ermöglichen, die später zum Gefühl von Präsenz beitragen können, welche das Gefühl beschreibt, sich wirklich in einer bestimmten Umgebung zu befinden (Mayer & Makransky, 2022). Dabei sind mögliche Konsequenzen für Lernprozesse bisher nur zu geringem Maße verstanden. Während vorhergehende Studien sich insbesondere auf kognitive Lerneffekte und Zusammenhänge mit einem erhöhten cognitive load bezogen, sollen im Vortrag Zusammenhänge zwischen dem Einsatz unterschiedlicher digitaler

Repräsentationsformen mit der Präsenz (F1), positiven Emotionen (F2) und der selbstberichteten kognitiven Aktivierung (F3) berichtet werden.

Um die aufgestellten Forschungsfragen zu beantworten wurde eine randomisierte quantitative Feldstudie mit 100 Lernenden (M = 15,69 Jahre) von zwei Schulen durchgeführt. Dabei wurden innerhalb der Klassenräume Stationen mit Texten, Tablets und VR-Brillen aufgebaut, auf die alle Lernenden gleichmäßig durch zufälliges Ziehen von Karten verteilt wurden. Nacheinander konnten die Lernenden dann immersive 360°-Videos entweder auf VR-Brillen (hohe Immersion) oder Tablets (mittlere Immersion) anschauen. Als weitere Vergleichsgruppe wurden ausgewählte Bilder sowie die sprachlichen Inhalte der 360°-Videos in Textform repräsentiert (gerine Immersion). Insgesamt wurden drei unterschiedliche ca. 4 minütige 360°-Videos eingesetzt, die sich auf die nachhaltige Landwirtschaft als ein Thema der Bildung für nachhaltige Entwicklung bezogen. Im Prä-test wurden dabei demographische Angaben und Einstellungen gegenüber der Landwirtschaft erhoben, welche sich nicht signifikant voneinander unterschieden. Nachdem die Messgüte mittels explorativer Faktorenanalysen (Ladungen > 0.40) und interner Konsistenzen überprüft und für gut befinden wurden, wurden die Forschungsfragen bearbeitet.

Dabei zeigte sich ein signifikanter Unterschied in der berichteten Präsenz ($F(2) = 38,06$, $p < 0.001$), wobei die höchste Präsenz in der VR-Gruppe berichtet wurde ($d = 0,70 - 2,30$). Es zeigten sich zudem signifikante Steigerungen der positiven Lernemotionen des Interesses ($F(2) = 24,36$, $p < 0.001$) und der Freude ($F(2) = 12,71$, $p < 0.001$). Auch hier berichteten die Lernenden die meisten positiven Emotionen für die VR, etwas geringere für Tablets und die geringsten für die Texte. Neben diesen Unterschieden innerhalb der Gruppen wurden Zusammenhänge zwischen der Präsenz und der Klarheit nach Wisniewski et al. (2020) als selbstberichtete Dimension der wahrgenomenen Unterrichtsqualität untersucht. Dabei zeigte war die Präsenz ein Mediator des Effektes der VR-Gruppe (codiert als 0 und 1) auf die Klarheit ($b = 0,23$, $p < 0.05$) mit Einstellungen als Kontrolvariable ($b = 0.41$, $p < 0.001$). Dieser Effekt ergab sich nur für die VR, aber nicht für die anderen Gruppen, weshalb nur für diese Gruppe von einem Zusammenhang von Einsatz eines Mediums, der Präsenz und der berichteten kognitiven Aktivierung ausgegangen werden muss.

Insgesamt ist die Studie eine Möglichkeit die Lerneffekte des Einsatzes immersiver VR besser zu verstehen, auch wenn die Rolle der selbstberichteten Daten kritisch diskutiert werden muss. Dabei erscheint es in den nächsten Jahren geboten nicht nur generelle Lerneffekte zu untersuchen, sondern diese stärker mit fachdidaktischen und bildungswissenschaftlichen Diskursen wie dem zur Unterrichtsqualität zu verknüpfen. Implikationen für weitere Studien und der Rolle von Unterrichtsqualität mit digitalen Medien werden abschließend diskutiert.

F-04: Bildungserfolg und -verlauf: Profile und Faktoren

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S06 S02 B06
Chair der Sitzung: Georg Karl Deutschmann

Geschlechtsspezifische Disparitäten in Bildung und Beruf. Untersuchung zu differenziellen Entwicklungsverläufen über die Lebensspanne.

Georg Karl Deutschmann[1], Michael Becker[1,2], Yi-Jhen Wu[1]
[1]Institut für Schulentwicklungsforschung (IFS) | TU Dortmund; [2]Abteilung Struktur und Steuerung des Bildungssystems, IPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Der individuelle Bildungs- und Lebensverlauf wird vielfältig vom Geschlecht mitgeprägt. Im Durchlaufen der Bildungsinstitutionen finden sich international und auch in Deutschland genderspezifische Disparitäten (Becker & McElvany, 2018; Stanat et al, 2018); Sowohl im Kompetenzerwerb wie auch in schulischen Leistungen sind weibliche Bildungsteilnehmende den männlichen Bildungsteilnehmenden mittlerweile relativ systematisch überlegen (Autorengruppe Bildungsberichterstattung, 2018; Hannover & Kessels, 2011). Dennoch verläuft der spätere berufliche Erfolg geschlechtsspezifisch differenziell. Entgegen den Disparitäten in der Bildung zeigen Arbeitnehmer entlang des Berufslebens, insbesondere hinsichtlich des Einkommens, günstigere Entwicklungen im Beruf als Arbeitnehmerinnen (Auspurg et al, 2017; Becker et al., 2019; Christofides et al., 2013). Trotz politischer Bemühungen zur Gleichstellung sind Frauen in hohen Berufspositionen signifikant unterrepräsentiert (Betrand, 2018; für Deutschland: Collischon, 2017) und in Sektoren mit geringem Einkommen überrepräsentiert (Brynin & Perales, 2015). Entsprechend steht zu klären, wo mit dem Einstieg in den Beruf und während des Berufsverlaufs diese Unterschiede zu verorten sind und wie sie sich verändern. Erklärungen fokussieren beispielsweise auf arbeitnehmerseitige Präferenzen oder arbeitgeberseitige Stereotypen gerade beim Berufseinstieg und der Berufswahl (Busch, 2013) oder längerfristig wirkende systemspezifische Benachteiligungen im Berufsleben (z. B. „glass ceiling effect“; Cotter et al., 2001; Grund, 2015).

Zudem sind die Disparitäten zwischen männlichen und weiblichen Arbeitnehmenden nicht homogen über Berufsfelder und -sektoren verteilt (Achatz, 2018; Holst et al., 2015) und es steht zu vermuten, dass dies auch durch Qualifikationen und Bildungsabschlüssen mitgeprägt wird. Gerade unterschiedliche Ausbildungsniveaus und -zertifikate sind sowohl für den Berufseinstieg als auch für die weitere berufliche Entwicklung zentral (Konietzka & Hensel, 2017). Inwieweit Geschlechterunterschiede im Beruf und vor allem eine geschlechterspezifische Entwicklungsdynamik von den Qualifikationen und Bildungsniveaus (mit-)geprägt sind, ist bislang noch nicht hinreichend geklärt; vor allem mangelt es an längsschnittlichen Perspektiven, die über die Veränderungen entlang der individuellen Lebensverläufe weiter Aufschluss geben.

Der Beitrag greift diese Fragen auf und untersucht geschlechterspezifische Unterschiede auf dem Arbeitsmarkt über die Zeit hinweg und in Abhängigkeit von Qualifkations- und Bildungsniveau, um zeit- und niveauspezifische Unterschiede in der beruflichen Entwicklung längsschnittlich zu beschreiben und zu erklären. Wir erwarten, dass ab dem Eintritt in den Arbeitsmarkt geschlechtsbezogene Unterschiede auftreten, die sich über die Zeit hinweg vergrößern. Für die geschlechtsspezifische berufliche Entwicklung wird ein differenzieller Effekt entlang der Bildung erwartet, wenngleich als offene Frage adressiert wird, wie stark diese Unterschiede über Bildungsstatus bzw. Berufsfelder differentiell über die Zeit hinweg ausfallen.

Die Untersuchung basiert auf der längsschnittlichen large-scale Studie *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter und jungen Erwachsenenalter (BIJU;* Baumert et al., 1996*).* Die Studie erfasst die Teilnehmenden von der siebten Klasse (1. Welle: 1991) bis ins Erwachsenenalter im Alter von knapp 40 Jahren (8. Welle: 2018) und ermöglicht es damit, Angaben aus der Jugendzeit mit den Entwicklungen im späteren Berufsleben in Verbindung zu bringen. Die hier zugrunde gelegte Stichprobe umfasst Angaben von n=4387 Personen. Die Angaben zur beruflichen Entwicklung umfassen drei Messzeitpunkte: den Berufseinstieg sowie das Alter von rund 30 Jahren und 40 Jahren.

Mithilfe latenter Regressionsmodelle (Cohen et al, 2010) wird querschnittlich und längsschnittlich getestet, welche messzeitpunktspezifische Geschlechterunterschiede auf Berufsstatus und Einkommen bestehen und inwiefern diese differentiell nach Bildungsstatus ausfallen (zudem unter Kontrolle sozialer und kognitiver Hintergrundmerkmale). Die berufliche Entwicklung wird dabei nach Berufsstatus und Einkommen getrennt betrachtet. Erste Analysen weisen auf eine differenzielle Entwicklung auf dem Arbeitsmarkt hin. Wir finden zum Berufseinstieg einen höheren durchschnittlichen Berufsstatus von Frauen. Einkommen und Berufsstatus der männlichen Personen sind zu späteren Zeitpunkten höher. Der Vortrag stellt weiterführende quer- und längsschnittliche Analysen vor. Theoretische und praktische Implikationen werden vertiefend diskutiert.

Do men and women differ when aspiring to an ideal occupation? Sex-separated as opposed to sex-combined samples in latent profile analyses

Divan Mouton, Florian G. Hartmann, Bernhard Ertl
Universität der Bundeswehr München, Germany

Students need to consider various factors before aspiring to and eventually pursuing a career. Holland's (1997) person-environment fit perspective would recommend an individual to ideally aspire to a career that is congruent with their vocational interests. Although Holland's recommendation is empirically supported (Nye et al., 2017), vocational interest also develops in relation to other contextual influences (Lent, Brown & Hackett, 1994).

Gottfredson's (1981; 2002) developmental theory of occupational aspiration posits that, in addition to interest congruence, both prestige and sextype also shape our career choice by eliminating occupations that we do not find acceptable. Sextype is a particularly central influence as it rules out a large proportion of contra-sextype occupations which, for example, can cause young girls to exclude highly prestigious, male-dominated careers .

Although several studies have investigated Gottfredson's career choice factors the results are mixed for various reasons (Junk & Armstrong, 2010). A person-centered approach may, however, illustrate varying career choice strategies which previous variable-centered approaches, that seek common mean structures across the sample population, could not express (Spurk et al., 2020).

We therefore posit that a latent profile analysis (LPA) could find distinct subtypes with different career choice profiles based on their interest congruence, prestige and sextype of their ideal aspiration. Furthermore, this study also investigates whether both sexes generate similar career choice profiles when analyzed within sex-separated samples.

This study therefore investigates interest congruence (Ertl & Hartmann, 2019), a standardized prestige (Christoph, 2005), as well as the same-sex proportions (SSP) of n=9277 German university starters' ideal occupational aspirations (NEPS; SC5:15.0.0; Blossfeld & Roßbach, 2019) as inputs to an LPA (MPlus v. 8.2, Muthén & Muthén, 2017). Each profile is labelled by their most prominent occupational aspirations, and each career choice dimension is categorized by high, typical and low levels. To investigate similar career choice strategies the latent profiles for the whole sample is cross-tabulated with both male (n=3598) and female (n=5679) latent profiles.

Eleven career choice profiles are presented as the best solution for the sex-combined sample. Although prestige produces the most variance amongst the eleven profiles, SSP highlight two contrasting, non-typical career choice profiles, namely *High Same-sex Aspirations* (P10, n=1565) in which female students aspire to become primary school teachers while males aspire to technical occupations, and *Low Same-sex Aspirations* (P4, n=414) in which males students aspire to work as primary school teachers while females aspire to technical occupations such as machine operating.

The sex-separated LPA's suggest 14 and 12 profiles as best solutions for males and females, respectively. The crosstabulations indicate that, although only one profile (P8 - *High Prestige Medical Aspirations*) is fully shared (100%) by both sexes, the majority of profiles cross-tabulate upwards of 75% of their sample onto the same profile, and up to 95% amongst 3/4 of female profiles. Further investigation of the career choice levels suggests that a greater number of male profiles (five, 46% of male sample) prioritize a higher same-sex aspiration (>70%) as compared to female profiles (one, 8% of female sample). In addition, smaller career choice profiles are expressed, such as the *High Same-Sex Computer Science Aspirations* (M14, n=282) and *High Same-Sex Aspirations* (M8, n=236) which include executive management and physics aspirations amongst males, and *Low Same-Sex Economic Aspirations (*F10, n=118) who aspire to executive management aspirations amongst females.

The current study found distinct career choice profiles amongst student‘s ideal aspirations, further supported by a majority of shared profiles from sex-separated LPA's. Unsurprisingly, differences in career choice strategies are present amongst the sexes, with males more inclined to aspire to high same-sex aspirations, while females aspire to relatively few low same-sex, and therefore male-dominated, occupations.

Fühlen sich geflüchtete Schülerinnen und Schüler in der Schule wohl? Eine Analyse des Einflusses von individuellen, sozialen und strukturellen Faktoren

Regina Becker, Gisela Will, Homuth Christoph, Horr Andreas
Leibniz-Institut für Bildungsverläufe, Bamberg, Deutschland

Bildungsbeteiligung und Bildungserfolg in Form von Noten und Abschlüssen sind zentrale Themen der Bildungsforschung. Zusätzlich und zunehmend rücken jedoch auch subjektive Aspekte schulischer Integration in den Vordergrund. So stellt subjektives Wohlbefinden in der Schule einen Wert an sich sowie eine kognitive und emotionale Grundlage für erfolgreiches Lernen und den Umgang mit schulischen Herausforderungen dar (z.B. Hascher & Hagenauer, 2011; Obermeier et al., 2021). Daher sollten für die Erklärung von Bildungserfolg subjektive Wahrnehmungen der Schülerinnen und Schüler stärker in den Blick genommen werden (z.B. Hascher et al., 2018).

Den Fokus auf subjektive Dimensionen von Bildungsteilhabe zu legen, ist insbesondere für geflüchtete Schülerinnen und Schüler zentral. Neben dem Erwerb von Kompetenzen und Bildungsabschlüssen, trägt der Schulbesuch durch Strukturierung des Alltags, welcher oftmals durch die Bedingungen im Herkunftsland oder die Flucht zerrüttet war (z.B. Matthews, 2008), auch maßgeblich zur (emotionalen) Stabilität und zur sozialen Integration von geflüchteten Kindern und Jugendlichen bei (z.B. McBrien, 2005; Pastoor, 2015). Als Quereinsteigerinnen und Quereinsteiger in ein nicht vertrautes Bildungssystem, mit oftmals geringeren Sprachkenntnissen in der Unterrichtssprache, möglicherweise nicht abgeschlossenen Asylverfahren oder dramatischen Erfahrungen während der Flucht sind geflüchtete Jugendliche zudem mit zusätzlichen Herausforderungen konfrontiert, ihren Bildungsweg zu meistern. Insgesamt wissen wir bislang relativ wenig über das schulische Wohlbefinden von jungen Geflüchteten in der Schule (für Ausnahmen siehe u.a. Pagel & Edele, 2021; Pastoor, 2015; Ziaian et al., 2018).

Wir nähern uns dieser Thematik an, indem wir das Wohlbefinden in der Schule von nach Deutschland geflüchteten Jugendlichen genauer untersuchen. *Wie fühlen sich geflüchtete Schülerinnen und Schüler in deutschen Schulen? Wovon hängt ihr subjektives Wohlbefinden ab?*

Bisherige Konzepte zur Erklärung subjektiven Wohlbefindens in der Schule unterscheiden Einflussfaktoren auf individueller (z.B. Geschlecht, Alter) und struktureller Ebene (z.B. Schultyp) (vgl. z.B. Hascher, 2004). Aber auch die soziale Dimension, wie etwa Unterrichtsqualität oder soziale Interaktionen mit Mitschülerinnen und Mitschülern werden zunehmend betont (Hascher & Hagenauer, 2011). Wir greifen auf diese theoretische Konzeption der empirischen Schulforschung zurück und erweitern diese um migrations- bzw. fluchtspezifische Faktoren. Individuelle Faktoren schließen daher auch mögliche posttraumatische Belastungsstörungen sowie zusätzliche Ungewissheiten durch unsichere Asylverfahren ein. Soziale Faktoren umfassen hier beispielsweise die Unterstützung von Geflüchteten durch Lehrkräfte. Hinsichtlich struktureller Faktoren berücksichtigen wir, dass unterschiedliche Beschulungsstrategien für Neuzugewanderte schulisches Wohlbefinden beeinflussen könnten (Pagel & Edele, 2021).

Unsere Analysen basieren auf Daten der ersten Welle der Panel-Studie „ReGES – Refugees in the German Educational System“ (Will et al., 2021). Die Stichprobe besteht aus n = 2,415 jugendlichen Geflüchteten, die zwischen 2014 und 2018 in Deutschland angekommen sind und zum ersten Interviewzeitpunkt die Sekundarstufe I des allgemeinbildenden Schulsystems besuchten. Wir analysieren die kognitive Dimension von Wohlbefinden (z.B. Diener et al., 2003) und operationalisieren subjektives Wohlbefinden über die Zustimmung zu der Aussage „Ich fühle mich wohl in der Schule“. Unsere Analysen basieren auf schrittweisen logistischen Regressionen mit robusten Standardfehlern. Fehlende Werte werden multiple imputiert (White et al., 2011).

Unsere Ergebnisse zeigen, dass sich jugendliche Geflüchtete insgesamt sehr wohl in der Schule fühlen. Fast drei Viertel stimmen der Aussage, dass sie sich in der Schule wohlfühlen, eher (28,1%) oder völlig (44,7%) zu. Multivariate Analysen ergeben, dass Faktoren, die aus der Schulforschung bekannt sind, auch das schulische Wohlbefinden von Geflüchteten beeinflussen (z.B. Schulleistungen, Gesundheit, soziale Eingebundenheit). Darüber hinaus sehen wir aber auch

signifikante Zusammenhänge zwischen dem schulischen Wohlbefinden und fluchtspezifischen Faktoren. So hängt etwa ein unsicherer Rechtsstatus negativ mit schulischem Wohlbefinden zusammen, während die wahrgenommene Unterstützung für Geflüchtete durch Lehrkräfte signifikant positiv mit dem Wohlbefinden in der Schule korreliert. Die Ergebnisse legen nahe, dass es wichtig ist, Lehrkräfte für die spezifische Situation von Geflüchteten zu sensibilisieren und das Unterstützungspotenzial durch die Lehrkräfte in den Schulen zu fördern.

Mental Health and Educational Attainment: How Developmental Stage Matters

Kaspar Burger[1,4], **Michael Becker**[2,3], **Ingrid Schoon**[4]
[1]Universität Zürich; [2]TU Dortmund; [3]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; [4]University College London

Theoretical background

Good mental health is a prerequisite for learning and educational performance (Auerbach et al., 2016). It helps students to pay attention in class, cope with difficulties in school, and make important educational transitions (Evans et al., 2018; Mojtabai et al., 2015). In contrast, poor mental health may disrupt educational trajectories and prompt early school leaving (Mikkonen et al., 2018; Slmoninski et al., 2011). However, mental health may differentially affect educational outcomes at different points in an educational career. Based on developmental science, it would be reasonable to expect individuals' mental health status to be more predictive of educational attainment during adolescence than in young adulthood. This is because adolescence is a dynamic maturational life phase that encompasses many developmental changes and challenges as well as potential turning points in human development (Dahl et al., 2018). Moreover, adolescents are more likely than young adults to lack the skills to cope effectively with mental health issues (Skinner & Zimmer-Gembeck, 2007). Another debate concerns whether mental health predicts educational attainment similarly among young men and women (Riglin et al., 2013). In fact, links between mental health and educational outcomes might vary by gender for various reasons—because girls and boys tend to experience substantially different types of mental health problems (Campbell et al., 2021), because there are gender-specific behaviors and responses to challenges in the educational domain (Eschenbeck et al., 2007), or because there are gender differences in how young people cope with common stressors and mental health problems (Wilhsson et al., 2017).

Research questions

With this in mind, we analyzed whether distinct dimensions of mental health predicted critical educational attainments differentially at distinct stages in an educational career and whether there were gender differences in the associations between mental health and educational attainment.

Method

We used data from Next Steps, a nationally representative panel survey of individuals born in 1989/1990 in England (N = 15,594, 48% female, 33% ethnic minority), to examine the extent to which distinct dimensions of mental health predicted males' and females' probabilities of (a) continuing full-time education beyond age 16, (b) enrolling at university by age 20, and (c) obtaining a university degree by age 25.

To assess mental health, we used the General Health Questionnaire (GHQ-12; Goldberg & Williams, 1988; see also Gnambs & Staufenbiel, 2018). While the GHQ-12 may be used to capture mental health in one sum score, research has also established its multidimensional nature (e.g., Padrón et al., 2012; Shevlin & Adamson, 2005).

We estimated logistic regression models to predict educational attainment at distinct stages in an educational career and we expressed the findings as odds ratios and as predicted probabilities for male and female respondents. We thus assessed the predictive power of different dimensions of adolescent and young adult mental health for three critical educational attainments (at ages 16, 20, and 25). We considered the complex sampling structure, applying survey weights (design weight, sample weight, and stratum weight) to correct for differential selection probabilities (Calderwood, 2018; DfE, 2011). To adjust parameter estimation in light of missing data, we used full information maximum likelihood (FIML) estimation.

Results and discussion

Results suggest that individuals' mental health was more consequential for educational attainment during adolescence than in young adulthood. Variation in specific dimensions of adolescent mental health significantly predicted later educational attainments, whereas among young adults, mental health was unrelated to educational attainment. Importantly, these findings did not vary by gender. This indicates a particular need to protect and foster adolescents' mental health to enable them to develop their academic abilities to their fullest potential.

F-05: Motivation, Selbstwirksamkeit, Selbstkonzept und Engagement

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S06 S03 B06
Chair der Sitzung: Laura Fink

Interesse und Anstrengung als Prädiktoren des Lernerfolgs

Laura Fink, Detlef Urhahne
Universität Passau, Deutschland

Interesse ist im Lernkontext ein wichtiges motivationales Konstrukt, welches mit intrinsischer Motivation und positiven Emotionen einhergeht (Hidi & Renninger, 2006). Es können verschiedene Arten des Interesses unterschieden werden: Situationales oder anfängliches Interesse vor Bearbeitung einer Lernaufgabe (Hidi & Renninger, 2006) und aktuelles Interesse – das momentane Erleben während der Bearbeitung einer Lernaufgabe (Knogler, 2017).

Bereits Dewey (1913) propagierte, dass Interesse gegenüber Anstrengung der Vorzug in der Bildung einzuräumen sei. Interesse als intrinsisch motivierter Lernzustand ist allerdings nur ein mäßiger Prädiktor für Lernleistungen (z.B. Lee & Stankov, 2018). Neuere Studien haben zudem gezeigt, dass Interesse während einer Lernaufgabe schnell abnehmen kann (z.B. Nuutila et al., 2021).

Anstrengung wird häufig als aversiv und als etwas beschrieben, das möglichst zu vermeiden sei, obwohl es einen Mehrwert schaffen kann (z.B. Inzlicht et al., 2018). Mit Anstrengung verbindet sich die Mobilisierung von Ressourcen während des Lernens (Gendolla & Wright, 2009). In neueren Metaanalysen hat sich Anstrengung als guter Prädiktor für Lernleistungen gezeigt (z.B. Silm et al., 2020).

Anstrengung kann zwischen sowie innerhalb von Aufgaben stark variieren (z.B. Heemskerk & Malmberg, 2020; Trautwein et al., 2015). Deshalb scheint es sinnvoll, die momentane Anstrengung zu messen, um Veränderungen während des Lernprozesses untersuchen zu können. Häufig wird Anstrengung nur als erlebte Anstrengung nach einer Aufgabe erfasst (z.B. Järvinen et al., 2022; Symonds et al., 2020). Wichtiger scheinen jedoch die Anstrengungsbereitschaft vor der Bearbeitung einer Lernaufgabe und die aktuelle Anstrengung während der Aufgabenbearbeitung zu sein, um Post-hoc Erklärungen zu vermeiden.

Die vorliegende Studie untersucht Interesse und Anstrengung als Prädiktoren des Lernerfolgs. Dabei liegt der Fokus auf dem Wechselspiel der beiden Konstrukte im Verlauf eines Lernprozesses. Es wird angenommen, dass Interesse und Anstrengung positiv mit dem Lernerfolg zusammenhängen. Lernprozesse sollten anfänglich intrinsisch motiviert sein und im weiteren Lernverlauf sollte Anstrengung an Bedeutung gewinnen.

An zwei experimentellen Laboruntersuchungen nahmen 152 Studierende (69.7% weiblich) im Alter von *M* = 22.24 Jahren (*SD* = 2.47) teil. Es wurden der Konzentrations-Leistungs-Test (KLT-R 6-13; Düker & Lienert, 2001) als eine Wissensanwendungsaufgabe und eine modifizierte Version des „Gedächtnis für Personen“-Tests (GPT; Pahlke & Bulla-Hellwig, 2002) als eine Wissenserwerbsaufgabe eingesetzt. Bei jedem Test wurde zu Beginn anfängliches Interesse und Anstrengungsbereitschaft sowie während der Testbearbeitung zu zwei Messzeitpunkten aktuelles Interesse und aktuelle Anstrengung mittels reliabler Fragebogenskalen (nach Gendolla & Wright, 2009; Hidi & Renninger, 2006) erfasst.

Mittels Korrelationen und linearen Strukturgleichungsmodellen wurden die angenommenen Zusammenhänge überprüft. Ein Vergleich der β-Koeffizienten im Strukturgleichungsmodell wurde mit einem Ansatz zur Neuparametrisierung von Kwan und Chan (2011) durchgeführt.

Beim KLT waren das anfängliche Interesse (β = .57, $p < .001$) und die aktuelle Anstrengung zu t_2 (β = .20, $p < .05$) signifikante Prädiktoren des Lernerfolgs. Beim GPT wurde nur die aktuelle Anstrengung zu t_2 (β = .64, $p < .05$) signifikant. Beim KLT wurde der Likelihood-Ratio-Test (LR-Test) hinsichtlich des anfänglichen Interesses und der Anstrengungsbereitschaft ($\Delta\chi2 = 19.905-0 = 19.905$, $\Delta df = 1-0 = 1$, $p < .05$) signifikant. Dies deutet darauf hin, dass der relative Effekt des anfänglichen Interesses auf die Lernleistung signifikant stärker war. Beim GPT wurde der LR-Test in Bezug auf das aktuelle Interesse und die aktuelle Anstrengung ($\Delta\chi2 = 3,829-0 = 3,829$, $\Delta df = 1-0 = 1$, $p = .05$) signifikant. Dies bedeutet, dass der relative Effekt der aktuellen Anstrengung auf die Lernleistung signifikant stärker ausfiel.

Interesse ist ein kurzzeitiger Motivationszustand, der in den Anfangsphasen des Lernens von Bedeutung ist. Im Lernverlauf führt Anstrengung zur Mobilisierung von Ressourcen und löst Interesse als Prädiktor der Lernleistung ab. Dieses Wechselspiel von Interesse und Anstrengung scheint auf verschiedene Arten von Lernprozessen übertragbar zu sein. Die Ergebnisse lassen vermuten, dass Lernleistungen durch Appelle an Anstrengung und Durchhaltevermögen im Lernprozess gesteigert werden können.

Beeinflussen Zielorientierungen und Selbstwirksamkeitserwartungen, wie Studierende tutorielles Feedback wahrnehmen und nutzen? - Ergebnisse einer Längsschnittstudie

Melanie V. Keller, Markus Dresel, Martin Daumiller
Universität Augsburg, Germany

Feedback spielt eine zentrale Rolle in Lehr-Lernkontexten und kann Lernen und Leisten positiv beeinflussen (Hattie & Timperley, 2007; Panadero & Lipnevich, 2022) – aber nur dann, wenn das Feedback von den Lernenden konstruktiv wahrgenommen und effektiv genutzt wird (Kollar & Fischer, 2010). Fragen nach der Wahrnehmung des Feedbacks durch Lernende sowie nach Faktoren auf Seiten der Lernenden, die zur tatsächlichen Nutzung führen, sind bisher unzureichend untersucht (Van der Kleij & Lipnevich, 2021).

Aus theoretischer Sicht sollte Motivation eine wichtige Rolle bei der Wahrnehmung und Nutzung von Feedback spielen, darunter insbesondere Zielorientierungen und Selbstwirksamkeitserwartungen (Dweck, 1986; Sherf & Morrison, 2019). Je nachdem, welche Ziele Lernende verfolgen, könnte die Implementierung von Feedback etwa als gute Lerngelegenheit (bei starken Lernzielen), als Gelegenheit, die eigenen Fertigkeiten zu präsentieren (starke Erscheinungsziele) oder als unwillkommener Aufwand (starke Arbeitsvermeidungsziele) wahrgenommen werden. Wie Lernende Feedback wahrnehmen/nutzen, sollte ebenso davon abhängen, inwiefern sie sich in der Lage sehen, auch bei auftretenden Schwierigkeiten erfolgreich mit Feedback umzugehen (hohe Selbstwirksamkeitserwartungen).

Wir nehmen an, dass diese motivationalen Aspekte mit der Wahrnehmung von Feedback (in Form von Leistungsemotionen sowie der Wahrnehmung des Feedbacks als hilfreich) sowie mit der quantitativen und qualitativen Nutzung von Feedback zusammenhängen. Dabei sollten hohe

Selbstwirksamkeitserwartungen, Annäherungslern- und Erscheinungsziele mit positiven Emotionen beim Lesen des Feedbacks, der Wahrnehmung als hilfreich und umfangreicherer Nutzung einhergehen; Vermeidungserscheinungsziele und Arbeitsvermeidungsziele sollten umgekehrte Effekte zeigen. Zudem ist davon auszugehen, dass Studierende Feedback eher nutzen, je mehr sie es als hilfreich wahrnehmen.

Zur Untersuchung dieser Zusammenhänge führten wir eine Feldstudie mit 204 Studierenden (133 weiblich; M_{Alter} = 20,77; SD_{Alter} = 2,49) in einem Hochschulkurs durch. Diese bearbeiteten wöchentliche Aufgabenblätter und erhielten darauf tutorielles Feedback, mit dem sie ihre Aufgabenblätter überarbeiten mussten. Wir erfassten, jeweils mit etablierten Skalen, die selbstbezogenen Ziele der Studierenden zu Beginn des Semesters und direkt nach dem Erhalt des Feedbacks ihre Selbstwirksamkeitserwartungen im Hinblick auf die Überarbeitung, erlebte Leistungsemotionen, ihre Einschätzung des Feedbacks als hilfreich und ihre Überarbeitungsintentionen. Nach der Überarbeitung schätzten die Studierenden ein, wie sehr sich die Qualität ihrer Ausarbeitung verbesserte, und wir erfassten das objektive Überarbeitungsausmaß durch computerbasierten Vergleich beider Abgaben (Mullen, 2020).

Strukturgleichungsmodelle bestätigten unsere Annahmen zu großen Teilen. Im ersten Messzeitpunkt hingen Selbstwirksamkeitserwartungen und Lernziele signifikant mit einer positiven Wahrnehmung des Feedbacks (Freude: β= .55/.23, Stolz: β=.56/.18, Wahrnehmung als hilfreich: β=.30/.29) und negativ mit einer negativen Wahrnehmung (Scham: β = –.44/–.09, Angst: β = –.60) zusammen. Annäherungserscheinungsziele zeigten einen signifikanten, positiven Zusammenhang mit Stolz (β = .13). Bezüglich der Nutzung des Feedbacks hatten Studierende mit höheren Annäherungsleistungszielen signifikant eher den Eindruck, dass sich die Qualität ihrer Ausarbeitung durch die Implementierung des Feedbacks verbessert hatte (β = .13), umgekehrt für hohe Vermeidungserscheinungsziele (β = –.10). Die Wahrnehmung des Feedbacks als hilfreich hing signifikant mit höheren Überarbeitungsintentionen (β = –.44) sowie einer höheren subjektiven Qualitätsverbesserung zusammen (β = –.53). Über das Semester hinweg blieben die Selbstwirksamkeitserwartungen mit der Wahrnehmung des jeweiligen Feedbacks assoziiert, die Bedeutung der Ziele nahm dagegen ab.

Insgesamt wurden die erwarteten, positiven Zusammenhänge von Selbstwirksamkeitserwartungen, Lernzielen und Annäherungserscheinungszielen größtenteils gefunden, jedoch waren Vermeidungserscheinungsziele und Arbeitsvermeidungsziele weniger relevant als erwartet. Während sich die fehlenden Effekte in späteren Messzeitpunkten dadurch erklären lassen, dass sich die zu Semesterbeginn angegebenen Ziele möglicherweise im Semesterverlauf änderten (Fryer & Elliot, 2007), bleiben die Gründe für die fehlenden Effekte im ersten Messzeitpunkt offen. Es zeigte sich jedoch, dass motivationale Aspekte, genauer: Selbstwirksamkeitserwartungen und selbstbezogene Ziele, einen Einfluss darauf haben, wie Studierende Feedback wahrnehmen und wie sie es nutzen. Es scheint sinnvoll, diese Zusammenhänge tiefergehend zu untersuchen. Aus praktischer Sicht scheint es zudem auch in Feedbackprozessen empfehlenswert, Lernziele zu fördern und Vermeidungserscheinungsziele zu reduzieren sowie sicherzustellen, dass die Lernenden Feedback als hilfreich wahrnehmen.

Schulengagement und Akkulturationsverhalten von Schülerinnen und Schülern während eines Auslandsschuljahres: Eine Längsschnittperspektive auf intra- und interindividuelle Dynamiken

Juan Serrano-Sánchez, Julia Zimmermann, Kathrin Jonkmann
FernUniversität in Hagen, Deutschland

Theoretischer Hintergrund:

In einer globalisierten Welt ist die internationale Bildungsmobilität von Lernenden ein wichtiges Instrument, um ihre Fähigkeit zum Agieren in modernen globalen Gesellschaften zu verbessern, mögliche kulturelle Barrieren zu überwinden und Diskriminierung zu bekämpfen (Knight, 2003). Die hier betrachtete Form internationaler Bildungsmobilität betrifft adoleszente Sojourners, d.h. Schüler*innen im Alter von 14 bis 18 Jahren, die ihr Heimatland für ein Schuljahr verlassen, um im Ausland die Schule zu besuchen. Zur Bewältigung der kulturellen Transition setzen die Schüler*innen individuelle Akkulturationsstrategien ein, indem sie ihr Verhalten in unterschiedlichem Maße am Gastland (host-cultural behavioral engagement) und dem Herkunftsland (home-cultural behavioral engagement) orientieren. Wie erfolgreich die Bewältigung der kulturellen Transition verlief, lässt sich an verschiedenen Indikatoren der Adaptation bemessen. Für die Schüler*innen während eines Auslandsschuljahres ist die Schule ein zentraler Bereich ihrer Auslandserfahrung, da sie dort einen substanziellen Teil ihrer Zeit verbringen und in Kontakt mit Mikrosystemen (z.B. Kommiliton*innen, Lehrkräften) und Makrosystemen (z.B. dem Schulsystem) der Gastkultur treten (Hein et al., 2017; Motti-Stefanidi & Master, 2013). Demnach ist das Schulengagement ein relevanter Indikator erfolgreicher Adaptation (Berry et al., 2006; Rienties et al., 2012).

Vor diesem Hintergrund betrachtete die vorliegende Studie die längsschnittlichen transaktionalen Zusammenhänge zwischen Schulengagement (affektiv, verhaltensbezogen und kognitiv) und Akkulturationsverhalten (host- und home-cultural behavioral engagement) von Schüler*innen während eines Auslandsschuljahres auf inter- und intraindividueller Ebene mithilfe eines random-intercept cross-lagged panel model (RI-CLPM, Hamaker et al., 2015) und trägt damit zum grundlegenden Verständnis der Akkulturations- und Adaptationsprozesse international mobiler Lernender bei.

Fragestellung:

Wie gestaltet sich die Beziehung zwischen Schulengagement und Akkulturationsverhalten adoleszenter Sojourners im Verlauf der Zeit auf intra- und interindividueller Ebene?

Methode:

Die Daten stammen aus den ersten vier Wellen des längsschnittlich angelegten Forschungsprojekts [Projektname]. Die Stichprobe umfasst deutsche Schüler*innen, die im Schuljahr 2016/17 einen Auslandsaufenthalt absolvierten. Alle Teilnehmenden, die mindestens einen der Fragebögen zwischen t0 und t3 ausgefüllt hatten (N = 1,299), wurden in die Analysen einbezogen. Soziodemographische und psychologische Merkmale wurden anhand etablierter Messinstrumente erhoben. Während des Auslandsaufenthalts wurde Schulengagement mittels der *Student Engagement in School Scale* (Lam et al., 2014; „Morgens freue ich mich meistens darauf, zur Schule zu gehen“) erfasst. Host- und home-cultural behavioral engagement wurden mit Bezug zu

verschiedenen relevanten Lebensbereichen der adoleszenten Sojourners erfasst (Autoren, 2021; „In den letzten vier Wochen habe ich mich mit Deutschen/Menschen aus meinem Gastland getroffen“).

Ergebnisse:

Das RI-CLPM zeigte eine sehr gute Anpassung an die Daten (χ^2 (20) = 30,415, p = .063, CFI = .998, RMSEA = .019, SRMR = .011). Auf der intraindividuellen Ebene zeigten die Zusammenhänge zwischen Schulengagement im Ausland und host- und home-cultural behavioral engagement, dass das affektive Schulengagement in beiden Messintervallen (t1-t2 und t2-t3) positiv mit dem späteren host-cultural behavioral engagement und negativ mit dem späteren home-cultural behavioral engagement zusammenhing. Ebenso wurde ein negativer Zusammenhang zwischen host-cultural behavioral engagement zu t2 und dem verhaltensbezogenen Schulengagement zu t3 identifiziert. Hingegen wurden keine signifikanten Zusammenhänge zwischen kognitivem Schulengagement und Akkulturationsverhalten auf der intraindividuellen Ebene identifiziert.

Auf der interindividuellen Ebene bestand eine signifikante positive Beziehung zwischen host-cultural behavioral engagement und allen drei Facetten des Schulengagements. Im Gegensatz dazu wurden keine signifikanten Zusammenhänge zwischen home-cultural behavioral engagement und den drei Facetten des Schulengagements identifiziert.

Diskussion:

Unsere Ergebnisse erweitern frühere Erkenntnisse, indem sie eine differenziertere Perspektive auf den (längsschnittlichen) Zusammenhang zwischen Schulengagement und Akkulturationsverhalten bieten. Unter anderem deuten sie darauf hin, dass die Förderung des affektiven schulischen Engagements ein guter Ansatzpunkt sein könnte, um das allgemeine host-cultural engagement zu unterstützen. Insbesondere könnten Schüler*innen von vorbereitenden Aktivitäten vor dem akademischen Jahr im Ausland profitieren, die ihr (affektives) Schulengagement erhöhen, wie z. B. Online-Interaktionen mit Gleichaltrigen und Lehrkräften im Ausland oder die Möglichkeit, ihre zukünftige Gastschule virtuell zu besuchen (Taylor & Parsons, 2011).

Effekte eines adaptiven Mathematik-Lernprogramms auf die Leistung, das Selbstkonzept und die Mathematikangst von Schüler:innen

Anna Hilz, Karen Aldrup
IPN - Leibniz Institute for Science and Mathematics Education, Germany

Mathematiklernprogramme sind in den letzten Jahrzehnten in den Fokus der Forschung gerückt, da sie durch Merkmale wie Adaptivität und Feedback dazu beitragen können, das Mathematiklernen zu verbessern (Hillmayr et al., 2020). Dennoch sind die Forschungsergebnisse zu Leistung heterogen und Gesamteffekte gering (d = 0.37; Byun & Joung, 2018). Pellegrini et al. (2021) stellen daher das anhaltende Forschungsinteresse an Mathematiklernprogrammen in Frage. Aber statt zu vermuten, dass Mathematiklernprogramme unwirksam sind, könnte man sich auch fragen, ob die Ansätze früherer Forschung in der Lage waren, potentielle Effekte aufzudecken. Studien erheben Leistung oft als Gesamtoutcome, statt zwischen den einzelnen Programminhalten zu differenzieren (z.B. Grundrechenarten; Ran et al., 2021). Außerdem ist wenig erforscht, wie sich das Übungsverhalten auf die Leistung auswirkt (Spitzer, 2021). Fehlende Effekte könnten auch dadurch zustande kommen,

dass nur die Bereitstellung von Lernprogrammen untersucht wird; Schüler:innen müssen ein Programm aber regelmäßig nutzen, um davon profitieren zu können. Darüber hinaus gibt es nur wenige Studien, die Effekte auf affektiv-motivationale Variablen berücksichtigen (Fadda et al., 2021).

Daher wurde in dieser Studie ein randomisiertes Prä-Post-Kontrollgruppendesign verwendet, um die Auswirkungen eines adaptiven computergestützten Arithmetiklernprogramms auf differenzierte Leistungsergebnisse (Addition und Subtraktion), Mathematikselbstkonzept und Mathematikangst zu überprüfen. Auch wurde das Lernverhalten in Form von geübten Aufgaben und geübten Wochen berücksichtigt.

RQ 1: Fördert ein solches Programm die Leistungen der Schüler:innen (Addition und Subtraktion), deren Mathematikselbstkonzept und eine Verringerung der Mathematikangst?

RQ 2: Beeinflusst das Übungsverhalten (Anzahl der geübten Aufgaben und Wochen) die Leistung der Schüler:innen (Addition und Subtraktion), deren Mathematikselbstkonzept und eine Verringerung der Mathematikangst?

Die Analysen basieren auf längsschnittlichen Daten von 357 Fünftklässler:innen (49% weiblich). 187 Schüler:innen wurden der Experimental- und 170 Schüler:innen der Kontrollbedingung zugewiesen. Zu Studienbeginn (T1, Herbst 2019) wurden folgende Zielvariablen am Tablet erhoben: Mathematikleistung mit dem Heidelberger Rechentest (Haffner et al., 2005), Mathematikselbstkonzept (Köller et al., 2000) und Mathematikangst (MARS 4-6; Roick et al., 2013). Danach erhielten die Schüler:innen der Experimentalgruppe Zugang zu dem adaptiven Lernprogramm Math Garden (Klinkenberg et al., 2011), welches überwiegend die Grundrechenarten trainiert. Das Lernverhalten wurde in Form von geübten Aufgaben und geübten Wochen erfasst (log files). Im Frühjahr 2020 (T2) wurden die genannten Variablen erneut erhoben.

Zur Beantwortung der Fragestellungen wurden Regressionsanalysen in Mplus gerechnet. Die Mehrebenenstruktur der Daten wurde mit „type=complex“ berücksichtigt. Für RQ1 betreffende Analysen war die UV die Gruppenzugehörigkeit und die AV jeweils die Additions- bzw. Subtraktionsleistung, das Mathematikselbstkonzept und die Mathematikangst. Zur Größeneinordnung signifikanter Effekte wurde zusätzlich Cohen's *d* berechnet. Für RQ2 wurde nur die Experimentalbedingung analysiert: Prädiktoren waren die Anzahl der geübten Aufgaben bzw. geübten Wochen auf die Zielvariablen. Fehlende Werte wurden über den Full Information Maximum Likelihood-Ansatz modellbasiert geschätzt.

In Bezug auf RQ1 stieg das Mathematikselbstkonzept in der Experimentalbedingung signifikant an ($\beta = .24$, $p = .003$; $d = 0.31$). In Bezug auf RQ2 konnte die Anzahl der geübten Aufgaben die Subtraktionsleistung ($\beta = .13$, $p = .001$), und die Anzahl der geübten Wochen die Additionsleistung vorhersagen ($\beta = .14$, $p = .030$).

Die Ergebnisse lassen vermuten, dass sich das Programm positiv auf das Mathematikselbstkonzept der Schüler:innen auswirkt. Effekte des Übungsverhaltens auf die Additions- und Subtraktionsleistung deuten darauf hin, dass das Lernverhalten von Schüler:innen relevant für die Leistungsverbesserung ist. Damit adressiert die Studie einen vielversprechenden Forschungsansatz: Statt der Gruppenzugehörigkeit sollte auch das Übungsverhalten berücksichtigt und die Leistung nach den im

Programm gelernten Inhalten differenziert werden. Es konnte kein Effekt auf die Mathematikangst gezeigt werden. Deskriptive Analysen zeigen, dass besonders mathematikängstliche Schüler:innen wenig mit dem Programm übten und folglich vermutlich nicht von dem adaptiven Algorithmus profitieren konnten. Zukünftige Forschung sollte sicherstellen, dass Schüler:innen und deren Lehrkräfte das Programm langfristig regelmäßiger nutzen.

F-06: Berufsschulunterricht gestalten

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S06 S04 B06
Chair der Sitzung: Kristina Schick

Informelles Lernen am Arbeitsplatz – Adaptierung und Validierung eines Instrumentes am Beispiel der ärztlichen Weiterbildung

Kristina Schick[1], Markus Heim[2], Christian M. Schulz[2], Frederick Schneider[2], Pascal O. Berberat[1], Martin Gartmeier[1]
[1]Technische Universität München, TUM School of Medicine, TUM Medical Education Center, Deutschland; [2]Technische Universität München, TUM School of Medicine, Klinik für Anästhesiologie und Intensivmedizin

Theoretischer Hintergrund

Informelles Lernen am Arbeitsplatz (Workplace Learning, WPL) ist eine wichtige Ressource für berufliche (Weiter-)Bildung und Kompetenzentwicklung (Eraut, 2004). Im Gegensatz zum formellen Lernen werden keine spezifischen Kompetenzen und Lernziele vordefiniert, die mittels des informellen WPL erworben werden sollen (Tynjälä, 2008). Somit stellt die Erfassung des Lernerfolg bzw. Kompetenzzuwachses durch informelles WPL eine Herausforderung dar. Kyndt et al. (2013) entwickelten ein Messinstrument zur Erfassung dieses Lernerfolgs basierend auf den drei Faktoren „generische Lernerfolge“, „berufsspezische Lernerfolge“ und „organisationsspezifische Lernerfolge“. Bisher wurde dieses Instrument nur im flämischen Raum bei Sozialarbeiter*innen und Polizist*innen angewendet (Janssens et al., 2016; Kyndt et al., 2013). Ziel unserer Studie ist es nun dieses Messinstrument für den deutschsprachigen Raum am Beispiel der ärztlichen Weiterbildung zu adaptieren und zu validieren. Zur Validierung des Instruments werden die ärztlichen Kompetenzfacetten „medizinische Expertise“, „Kommunikation“ und „Wissenschaftlichkeit“ herangezogen, die wichtige Aspekte des ärztlichen Profession abbilden (Frank et al., 2015).

Fragestellung

Die Studie untersucht die folgenden Forschungsfragen:

(1) Weist die adaptierte Version des Fragebogens zum informellen WPL gute psychometrische Eigenschaften auf (Reliabilität: $\Omega \leq .7$; Konstruktvalidität: Exploratorische Faktorenanalyse mit guten Modellfitwerten der Messmodelle und des Strukturgleichungsmodell)?

(2) Welchen Einfluss haben die Selbsteinschätzung der Kompetenzfacetten „Medizinische Expertise“, „Kommunikation“ und „Wissenschaft“ auf die Lernerfolge des informellen WPL?

Methode

Um die Fragestellung zu beantworten, nutzten wir Daten des Bayerischen Absolventenpanels Medizin (Medi-BAS) im Erhebungsjahr 2018/19. Die Studie inkludierte 528 Ärzt:innen in Weiterbildung (n = 339 weiblich; Alter: M = 29,79; SD = 3,37 Jahre), die an einer bayerischen medizinischen Fakultät im Wintersemester 2016 oder Sommersemester 2017 approbiert worden waren. Die Teilnehmer*innen beantworteten den für den deutschsprachigen Raum und die ärztliche Weiterbildung adaptierten Fragebogen zum informellen WPL (Kyndt et al., 2013) und den Freiburger Fragebogen zur Erfassung von Kompetenzen in der Medizin (FMK, Giesler et al., 2011). Mithilfe einer explorativen Faktorenanalyse wurde die Faktorenstruktur des Fragebogens für informelles WPL untersucht. Als Indikator der Reliabilität des Instruments wurde McDonald’s Omega berechnet. Zusammenhänge zwischen den WPL-Faktoren und den FKM-Kompetenzfacetten wurden mittels des Spearman’s rho Korrelationskoeffizienten berechnet. Zur Messung der Konstruktvalidität des WPL-Fragebogens

wurde ein Regressionsmodell mit den FMK-Kompetenzfacetten als Prädiktoren berechnet. Ein Signifikanz-Niveau von $p \leq .05$ wurde angenommen.

Ergebnisse und ihre Bedeutung

Die Ergebnisse der Forschungsfrage 1 zeigen eine gute Passung eines vierfaktoriellen Modells (χ^2 = 380,69; *df* = 113; CFI = 0,924; RMSEA = 0,071; SRMR = 0,048). Die vier WPL-Faktoren weisen jeweils gute Reliabilitäten ($\Omega \geq .69$) auf. Folgende Aspekte des informellen WPL in der ärztlichen Weiterbildung können beschrieben werden: Faktor 1 – Generische Lernerfolge: Kompetenzentwicklung, Faktor 2 – Generische Lernerfolge: Reflexion; Faktor 3 – Berufsbezogene Lernerfolge: Patientenfokus; 4 – Organisationsbezogene Lernerfolge. Somit können die drei Aspekte, die von Kyndt et al. (2013) postuliert wurden (Generische, berufsbezogene, organisationsbezogene Lernerfolge) ebenfalls in der deutschsprachigen Adaption des Messinstruments identifiziert werden.

Hinsichtlich Forschungsfrage 2 zeigen sich statistische signifikante Zusammenhänge zwischen den vier WPL-Faktoren und den drei FMK-Kompetenzfacetten „Medizinische Expertise“, „Kommunikation“ und „Wissenschaft“ (r = ,20 bis r = ,38). Die Ergebnisse des Strukturgleichungsmodells zeigten, dass die Kompetenzfacette „Medizinische Expertise“ einen Einfluss auf die vier WPL-Faktoren hat. Desweiteren scheint die Kompetenzfacette „Kommunikation“ auf die WPL-Faktoren 1 und 3 und die Kompetenzfacette „Wissenschaft“ auf die WPL-Faktoren 1 und 2 zu wirken.

Die Ergebnisse der Studie geben Grund zur Annahme, dass das Messinstrument sinnvoll verwertbare Informationen über Ergebnisse von informellem WPL während der ärztlichen Weiterbildung liefern kann.

„Wo die wilden Kerle wohnen“? Klassenkontext, Klassenkomposition und Klassenführung an beruflichen Schulen: eine kontextorientierte (Mehrebenen)Analyse

Sylvia Rahn, Christoph Fuhrmann, Inga Allkemper, Anna Shevchuk, Sabine Glock
Bergische Universität Wuppertal, Germany

Hintergrund und Fragestellung

Die Qualität der Klassenführung ist nachweislich von hoher Bedeutung für die fachlichen Schülerleistungen (Helmke, 2017). Dabei profitieren offenbar Schüler:innen mit Migrationshintergrund besonders (Seiz et al., 2016).
In der Unterrichtsforschung wird die Klassenführung als eine Basisdimension des Unterrichts begriffen, mit der die Lehrperson simultan die Schülerleistungen beeinflusst und günstige Voraussetzungen für andere Prozessmerkmale der Unterrichtsqualität schafft (Marsh et al., 2012). Eine wirksame Klassenführung, manifestiert sich in einer störungsarmen Lernumgebung mit einer hohen durchschnittlichen aktiven Lernzeit und wird durch präventive, proaktiven und reaktive Aspekte des Umgangs mit Unterrichtsstörungen und des Managements der Unterrichtszeit begünstigt (Helmke & Helmke, 2014). Zudem spielt für eine wirksame Klassenführung eine positive und freundliche Kommunikation (Mägdefrau, 2010) sowie ein unterstützendes und verständnisvolles Führungsverhalten der Lehrkraft (Wubbels & Brekelmans, 2005) eine Rolle.

In der jüngeren Unterrichtsforschung stellen „vermittelte Prozess-Produkt-Modelle“ der Unterrichtsqualität eine rein lehrerzentrierte Forschungsperspektive auf die Klassenführung in Frage.

Deshalb werden zunehmend auch die Bedeutung des Klassenkontextes und der Klassenkomposition für die Klassenführung und ihre Wirksamkeit untersucht (Fauth et al 2020; Göllner et al. 2020; Praetorius et al., 2016).

Vor allem für die Klassenführung an beruflichen Schulen sind Kontexteffekte und die Bedeutung der Klassenkomposition für die Klassenführung wahrscheinlich, aber noch kaum untersucht: Lehrkräfte an beruflichen Schulen sind hier nicht nur an einer Schulform tätig. Sie unterrichten stattdessen in einer schulischen Organisation, die in allen Bundesländern mehrere Schulformen und noch mehr Bildungsgänge umfasst. Zudem sind allein die Fachklassen des dualen Systems an sich bereits sehr unterschiedlich zusammengesetzt.

Nun gelten in der nationalen wie internationalen Literatur zur Klassenführung kulturell und ethnisch heterogen zusammengesetzte Klassen in großstädtischen Schulkontexten sowie Klassen mit einem hohen Anteil männlicher Lernender als besonders herausfordernd (Downey & Pribesh, 2004; Schönbächler et al., 2011). In der berufs- und wirtschaftspädagogischen Diskussion wird gar unter dem plakativen Titel „Wo die wilden Kerle wohnen“ von einem „Jungenproblem in Klassen des Übergangssystems“ gesprochen (Sloane, 2014).

Vor diesem Hintergrund wird im Vortrag untersucht, inwieweit das Kommunikations- und Führungsverhalten der Lehrperson sowie die Effektivität der Klassenführung an beruflichen Schulen mit dem Bildungsgang als Kontextmerkmal und dem Anteil männlicher Lernender sowie der Quote von Schüler:innen mit Migrationshintergrund als Kompositionsmerkmale zusammenhängen.

Sample und Datenqualität

Die Stichprobe wurde im Sommer 2022 im Rahmen eines laufenden DFG-projekts an 24 berufsbildenden Schulen gewonnen und umfasst über 3200 Schüler*innen aus 251 Klassen.

Die Klassenführung und ihre Effektivität wurden im Anschluss an das Wirkungsmodell der Klassenführung sensu Helmke & Helmke (2014) operationalisiert und das Kommunikations- und Führungsverhalten mit einer gekürzten Variante des Questionnaire on Teacher Interaction (QTI) gemessen.

Die Prozessmerkmale der Klassenführung und die „aktive Lernzeit“ und „Störungsfreiheit“ im Unterricht als Erfolgskriterien des Klassenführung wurden mit in der Regel zufriedenstellenden bis sehr guten internen Konsistenzen aus Sicht der Schüler:innen und Lehrkräfte erhoben.

Die Intrakassenkorrelationen induzieren für beide Erfolgskriterien der Klassenführung, dass eine simultane regressionsanalytische Auswertung der Daten auf zwei Ebenen notwendig ist (Störungsfreiheit: ICC1 = 0,31; aktive Lernzeit: ICC2 = 0,21).

Deren Ergebnisse werden im Vortrag präsentiert.

Befunde

Erste – noch auf Individual- und auf Klassenebene getrennte – Analysen deuten bei den individuellen Schülerratings auf schwache Effekte des Geschlechts und des besuchten Bildungsgangs auf die Störungsfreiheit des Unterrichts und (in noch geringerem Maße) auf die Nutzung der Lernzeit hin.

Für die aggregierten Ratings der Schüler:innen zeigen sich etwas stärkere Zusammenhänge. Tatsächlich sinken die Störungsfreiheit und die aktiv genutzte Lernzeit tendenziell mit zunehmendem

Männer- und Migrantenanteil in den Klassen. Teils verschwinden diese Effekte, wenn der besuchte Bildungsgang kontrolliert wird.
Insgesamt deuten die Ergebnisse darauf hin, dass die Klassen des Übergangssegments und Fachklassen des dualen Systems in Männerberufen etwas schwerer zu führen sind als andere Lerngruppen der beruflichen Schulen. Der Vortrag mündet in die Diskussion der praktischen Implikationen der Befunde.

Macro-Scaffolding im Berufsschulunterricht – Analyse von Kohärenz und Adaptivität als Maße für die Qualität von Unterrichtsentwürfen

Benjamin Herbert, Manon Heuer-Kinscher, Rico Hermkes, Minnameier Gerhard
Goethe Universität Frankfurt am Main, Deutschland

Theoretischer Hintergrund

Für gelingenden Unterricht und das Erreichen einer hohen Unterrichtsqualität ist u.a. die Passung zwischen Instruktion und Lernprozessen von Bedeutung. Das beginnt nicht erst im Unterrichtsgeschehen, sondern bereits bei der Planung und den dabei vorgenommenen Strukturierungsmaßnahmen. Diese sollen für die Lernenden in einer Weise passend sein, dass sie einerseits potentielle Lernhürden antizipieren und intendierte Lernprozesse durch Hilfestellungen erleichtern, andererseits aber nicht die Eigenaktivität der Lernenden zu stark einschränken. Diese Art von Unterstützung wird als Scaffolding bezeichnet (Hermkes et al., 2018; van de Pol & Elbers, 2013). Dabei wird zwischen Macro- und Micro-Scaffolding differenziert, wobei das im Fokus unseres Beitrags stehende Macro-Scaffolding die Strukturierung und Sequenzierung des Unterrichts sowie die in Unterrichtsmaterialien implementierten Lernhilfen betrifft.

Die Bestimmung der Qualität von Unterrichtsplanungen erfolgt bislang in der Regel mittels Ratings, die in den Unterrichtsentwürfen explizierte Einzel-Maßnahmen – z.B. solche zur Binnendifferenzierung – fokussieren, aber nicht systematisch den geplanten Unterricht als Prozess rekonstruieren (vgl. hierzu insb. die Arbeiten von König et al., 2020, zu didaktischer Adaptivität von Unterrichtsplanungen). Das kann zu Validitätsproblemen führen. Demgegenüber schlagen wir mit „Kohärenz“ und „Adaptivität“ zwei systematisch differenzierte Kriterien vor, für die wir entsprechende Maße entwickelt haben, die sich auf den Prozess beziehen. Kohärenz bezieht sich dabei auf die logische Folgerichtigkeit (im Sinne einer umfassenden Erkenntnislogik) und Vollständigkeit der einzelnen Erarbeitungsschritte, wohingegen Adaptivität eben die Passung zu den Unterstützungsbedarfen der Lernenden im o. g. Sinn beschreibt.

Fragestellung

In unserer Studie untersuchen wir, wie Lehrkräfte ihren Unterricht planen und dabei Macro-Scaffolds implementieren. Dabei steht die Entwicklung einer Erhebungsmethode, d.h. die Erfassung von Scaffolding-Mustern und deren Kohärenz sowie Adaptivität, im Vordergrund. Im Anschluss daran analysieren wir, inwieweit die identifizierten Macro-Scaffolding-Muster bestimmte Prozessqualitäten im Unterrichtsgeschehen, wie auftretenden Lernschwierigkeiten, bedingen.

Methode

Untersucht werden 30 Klassen an beruflichen Schulen in Hessen. Der Unterricht umfasst eine 90-min. Einheit zum Thema „Quantitativer und qualitativer Angebotsvergleich“, gehalten von der jeweiligen Fachlehrperson. Datengrundlage sind die in Drehbuchform ausgearbeiteten Unterrichtentwürfe, der videographierte Unterricht sowie die Tablet-Streams in den Einzel- und Gruppenarbeitsphasen der SuS. Kodiert werden Lernsequenzen hinsichtlich ihrer inferentiellen Charakteristik, dem Inhaltsbereich sowie der Wissenskonstruktionstätigkeit der SuS (eigenständig konstruiert, im Plenum ko-konstruiert oder nachvollzogen).

Ergebnisse und ihre Bedeutung

Ein Kodiersystem wurde entwickelt und für eine Teilstichprobe von 15 Unterrichtsentwürfen (insg. 140 Unterrichtssequenzen) erprobt. Die Raterübereinstimmung (Cohen‘s Kappa) lag für die Kodierungen der Inferenzen/Lernschritte im Bereich zwischen K=.61 - .72 und für die Wissenskonstruktionsaktivität (selbst konstruiert, ko-konstruiert, nachvollzogen) bei K=.88. Das Kodiersystem wurde anschließend überarbeitet und wird aktuell für die Kodierung weiterer Unterrichtsentwürfe eingesetzt.

Vorgestellt werden das finale Kodierverfahren sowie aktuelle Ergebnisse, die über Kohärenz und Adaptivität der Unterrichtsplanungen Aufschluss geben. In den bisher analysierten Daten zeigen sich typische Muster mangelnder Adaptivität bzw. von Inkohärenz, die Lernschwierigkeiten bei den SuS und einen erhöhten Bedarf an Micro-Scaffolding im Unterrichtsgeschehen erwartbar machen. So zeigt sich, dass Lernschritte in der Planung ausgelassen werden, die für die logische Verknüpfung von Inhaltsbereichen maßgeblich sind und dass essenzielle Lerninhalte (die u.a. den systematischen Aufbau der Angebotsvergleichsschemata betreffen) durch die Lehrkraft oder die Unterrichtsmaterialien vorgegeben werden und von den Lernenden nur rezeptiv verarbeitet werden. Eine aktiv-konstruktive Auseinandersetzung erfolgt dagegen selten.

Zum einen können die Ergebnisse einen Beitrag zur Unterrichtsqualitätsforschung leisten und als Prädiktoren in statistische Analysen einfließen, um Lernunterschiede zwischen den SuS bzw. zwischen den Klassen aufzuklären. Zum anderen können sie auch im Rahmen der Aus- und Weiterbildung von Lehrkräften genutzt werden, um so unterrichtliche Planungskompetenzen zu fördern.

Virtual Reality und kognitive Aktivierung in der beruflichen Bildung: Ergebnisse einer quasi-experimentellen Studie bei Auszubildenden im Einzelhandel

David Kablitz, Matthias Conrad, Stephan Schumann
Universität Konstanz, Germany

Theoretischer Hintergrund

Mit dem Einsatz von Virtual Reality (VR) gehen neuartige Möglichkeiten zur Gestaltung von Lehr-Lernumgebungen einher, die sich auch im Bereich der beruflichen Bildung niederschlagen, insbesondere im Hinblick auf die Simulation berufstypischer Arbeitssituationen (Hellriegel & Čubela, 2018). VR-Technologien werden bislang überwiegend im technisch-gewerblichen Bereich eingesetzt, für die kaufmännische Berufsbildung liegen bis dato kaum belastbare Befunde über Effekte der VR-Nutzung vor (Hamilton et al., 2020; Conrad et al., im Druck). Gleichzeitig stellt gerade die didaktische Gestaltung eines VR-basierten Unterrichts und dadurch auch die Sicherstellung der

Unterrichtsqualität eine Herausforderung für Lehrkräfte dar (Hellriegel & Cubela, 2018). Insbesondere im Hinblick auf die Tiefenstrukturen (kognitive Aktivierung, Classroom-Management, konstruktive Unterstützung) werden somit an die Lehrkräfte neue Herausforderungen gestellt, um einen möglichst lernförderlichen Unterricht zu gestalten.

Obwohl in der Literatur auf die Relevanz einer sinnvollen didaktischen Einbettung von VR hingewiesen wird (Hellriegel & Cubela, 2018; Buchner & Aretz, 2020), ist bislang wenig darüber bekannt, wie ein gelingender Lehr-Lernprozess mit VR-Technologie im Hinblick auf die Tiefenstrukturen des Unterrichts gestaltet werden sollte (Hellriegel & Cubela, 2018).

In der Literatur wird in dieser Hinsicht darauf hingewiesen, dass die Effektivität von digitalen Medien im Unterricht maßgeblich von der kognitiven Aktivierung der Lernenden abhängt (Quast et al. 2021). Speziell die Möglichkeit eines handlungsbezogenen und somit anwendungsorientierten Unterricht stellt einen wesentlichen Vorteil von VR im Vergleich zur anderen Medien dar (Hellriegel & Cubela, 2018; Conrad et al., im Druck). Die Frage, inwiefern dadurch die kognitive Aktivierung mittels VR-Technologie im Unterricht gefördert werden kann, bleibt aber bislang weitestgehend ungeklärt.

Der vorliegende Beitrag thematisiert hierzu erste Erkenntnisse, welche anhand einer explorativen Studie im kaufmännischen Bildungsbereich gewonnen werden konnten.

Fragestellungen

- Inwiefern kann durch VR-basierten Unterricht die kognitive Aktivierung bei Schüler*innen gefördert werden?
- Inwiefern steht die kognitive Aktivierung in Zusammenhang mit dem Lernerfolg in VR-basiertem Unterricht?

Methode

Auf Basis einer explorativen quasi-experimentellen Feldstudie mit Pre-, Prozess- Post-Test-Design wurde eine VR-basierte Unterrichtsserie mit kaufmännischen Auszubildenden im Fach *Betriebswirtschaft* durchgeführt. Durch das verwendete Experimental-Kontrollgruppen-Design wurde ein Vergleich zu einem nicht-VR-basierten Unterricht ermöglicht. Hierdurch konnten neben dem themenbezogenen Wissenserwerb zum Thema *Verkaufsraumgestaltung* auch zentrale Aspekte der Tiefenstrukturen des Unterrichts und somit der kognitiven Aktivierung adressiert werden. Die verwendete virtuelle Umgebung in Gestalt eines virtuellen Supermarktes wurde zu diesem Zweck erstmalig in Bildungskontext eingesetzt und untersucht. Als Stichprobe werden vier Klassen im Ausbildungsberuf *Verkäufer*in* (1. Lehrjahr) herangezogen (N=79).

Ergebnisse und ihre Bedeutung

Die Ergebnisse weisen darauf hin, dass durch den Einsatz von VR-Technologie die kognitive Aktivierung bei Lernenden gefördert werden kann. Im vorliegenden Beitrag können signifikante Unterschiede in der kognitiven Aktivierung zwischen Experimental- und Kontroll-Gruppe (d=.68, p=.02) berichtet werden, die eine Vorteilhaftigkeit von VR-basiertem Unterricht hinweisen. Zudem kann ein signifikanter Zusammenhang zwischen der Leistung im eingesetzten Fachtest und der kognitiven Aktivierung festgestellt werden.

Die Studie liefert insgesamt erste Erkenntnisse zum VR-basierten Lehren und Lernen in der kaufmännischen Berufsausbildung. Mit Blick auf die Bildungspraxis werden hierdurch wichtige

Hinweise über Gelingensbedingungen von VR-gestützten Unterrichtssettings gewonnen und gleichzeitig Erkenntnisse für künftige Forschungsvorhaben erzielt.

F-07: Teilhabe und deren Ermöglichung

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S05 T00 B59
Chair der Sitzung: Sebastian Ixmeier

Teilhabe revisited – Kompetenzorientierte Aktivierung als Polarisierungseffekt in der Arbeitsverwaltung

Sebastian Ixmeier
Universität Duisburg-Essen, Deutschland

Die politisch-normative Grundlage der bundesdeutschen Arbeitsverwaltung stellt seit rund zwei Jahrzehnten das Aktivierungsparadigma dar. Dessen zentrale Implikationen suggerieren eine simultane Ausrichtung des Vermittlungshandelns entlang der Leitprinzipien Fördern und Fordern (Freier 2016). Nicht zuletzt aufgrund einer zunehmenden Verstetigung von Langzeitarbeitslosigkeit steht dieses zentrale Organisationsprinzip der Arbeitsverwaltung zunehmend in der Kritik (Brussig 2019). Entsprechend gewinnt aktuell ein neues Leitbild an Relevanz, das einer allgemeinen Förderung von Teilhabe mehr Aufmerksamkeit schenkt. Die Akteur*innen der beruflichen Rehabilitation sind somit zwei konträren Handlungslogiken ausgesetzt, die zwischen beruflichem Leistungsgedanken und allgemeinen „Entwicklungsmöglichkeiten in Form von Kompetenzen“ (Biermann 2008: 121) liegen. Anhand von aktuellen Befunden aus einem Forschungsprojekt zur beruflichen Rehabilitation von langzeitarbeitslosen Personen mit gesundheitlichen Einschränkungen, sollen die Deutungsschemata der Akteur*innen innerhalb dieses Spannungsfeldes dargestellt, die Aushandlungsprozesse nachgezeichnet und die Handlungslogiken und Folgen kritisch erörtert werden.

Die empirische Datengrundlage entstammt der wissenschaftlichen Begleitforschung des *rehapro*-Modellprojektes *Essen.Pro.Teilhabe*. Dabei arbeiten in einem interdisziplinären Kontext verschiedene Träger (u.a. Ärzt*innen, Psycholog*innen, Sozialarbeitende, Vermittlungsfachkräfte), das Jobcenter sowie potenzielle Arbeitgebende zusammen, um eine bestmögliche, ganzheitliche Betreuung der erwerbsfernen Zielgruppe mit (drohenden) Rehabilitationsbedarf zu ermöglichen. Der methodische Ansatz ist ein *Mixed-Methods-Untersuchungsdesign* mit einer Kombination aus qualitativen und quantitativen Forschungsmethoden. Die Datenbasis für die Forschungsarbeit bilden Online-Befragungen sowie leitfadengestützte Interviews mit Teilnehmenden, Projektmitarbeitenden sowie Stakeholder*innen (u.a. Arbeitgebende und Bildungsträger*innen). Darüber hinaus kommen fokussierte ethnographische Beobachtungsverfahren sowie ein prozessbegleitendes Monitoring, basierend auf Projektverlaufsdaten, zum Einsatz. Der so akquirierte, qualitative Datenkorpus wird fortlaufend mittels inhaltsanalytischen sowie rekonstruktiven Verfahren ausgewertet. Bei der Analyse des quantitativen Datenmaterials kommen deskriptive sowie inferenzstatistische Methoden zum Einsatz.

Der Beitrag zeichnet mit der kompetenzorientierten Aktivierung die Entstehung und Umsetzung eines hybriden Leitbildes in der Arbeitsverwaltung kritisch nach. Im Zentrum des Verwaltungshandelns steht nun mehr nicht nur eine berufliche Teilhabe, sondern ein inkrementelles, ganzheitliches Vorgehen in welcher die „Persönlichkeit [der Teilnehmenden] zum Orientierungspunkt beruflichen Handelns“ (Epping et al. 2001: 152 [Verändert durch SI]) wird. (Schnelle) Erwerbsintegration stellt demnach nicht mehr die alleinige Handlungslogik der Akteur*innen der Arbeitsverwaltung da, die sich innerhalb dieses unsicheren Spannungsfeldes bewegen müssen. Für die Betroffenen zeigen sich positive

Effekte mit Blick auf den individuellen Kompetenzerwerb. Allerdings scheint die erfolgreiche Umsetzung voraussetzungsvoll mit Blick auf deren Möglichkeitsspielräume sowie Veränderungsbereitschaft. Sind diese Parameter unzureichend ausgeprägt, drohen Exklusionsprozesse.

Messung der wahrgenommenen Person-Umwelt-Passung – welches Konstrukt wird wirklich erfasst?

Brit Junkuhn, Julian Etzel, Gabriel Nagy
Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Kiel, Germany

Theoretischer Hintergrund:

Die Person-Umwelt-Passung (PUP) beschreibt, wie gut die Merkmale einer Person zu korrespondierenden Merkmalen einer (beruflichen) Umgebung passen (Kristof-Brown et al., 2005). Theoriegemäß führt eine hohe PUP zu wünschenswerten Outcomes (z.B. Zufriedenheit, Leistung). Häufig wird die subjektiv wahrgenommene PUP direkt mithilfe von Selbstauskunftsfragebögen erfasst (direkte PUP). Direkte PUP-Messungen scheinen tatsächlich stark mit beruflichen Outcomes zu korrelieren, wobei hinterfragt wird, ob es sich um einen kausalen Effekt handelt. So erwägen einige Autoren, dass direkte PUP-Maße und Outcomes (z.B. Zufriedenheit) konfundiert sind und dieselben affektiven Reaktionen auf die berufliche Umgebung widerspiegeln (Edwards et al., 2006). Diese Sichtweise wird durch Tagebuchstudien gestützt, in denen die direkte PUP keinen Effekt auf spätere Erhebungen der Outcomes aufweist (Gabriel et al., 2014; Vleugels et al., 2018). Hier gilt jedoch zu beachten, dass die Ergebnisse von Cross-Lagged-Effekten nicht zwangsläufig tatsächliche Ursache-Wirkungs-Beziehungen aufdecken (z.B., Driver, 2022).

Fragestellung:

Die Studie untersucht, ob unterschiedliche direkte PUP-Maße im Sinne einer globalen Umgebungseinschätzung interpretiert werden können. Unsere Analysen setzen an der Multidimensionalität der direkten PUP an (z.B. Etzel & Nagy, 2021), wonach die unterschiedlichen direkten PUP-Dimensionen differenzielle, theoretisch plausible Zusammenhänge mit objektiven Prädiktoren aufweisen sollten. Hohe Korrelationen der direkten PUP-Dimensionen untereinander sowie ähnliche Beziehungen zu theoretisch postulierten objektiven Prädiktoren würden hingegen für die Annahme einer globalen Umgebungsevaluation sprechen.

Methode:

Unsere Analysen basieren auf einer großen, repräsentativen Stichprobe mit 2495 Probanden (35.1% weiblich) aus drei verschiedenen Ausbildungsberufen, die nach der Hälfte ihrer Ausbildungszeit befragt wurden (ManKobE; Retelsdorf et al., 2013). Mittels Selbstauskunftsfragebögen wurden vier direkte PUP-Dimensionen mit jeweils drei Items erfasst („Interessen-Tätigkeiten-Passung“ (Schmitt et al. 2008; α = .62), „Fähigkeiten-Anforderungen-Passung“ (Cable & DeRue, 2002; α = .73), „Bedürfnis-Angebot-Passung“ (Cable & DeRue, 2002; α = .87), „Person-Organisation-Passung“ (Cable & DeRue, 2002; α = .88)). Zudem ließ sich aus den separaten Informationen bezüglich personenseitiger Interessen und umweltseitiger Tätigkeitseinschätzungen ein objektiveres Interessen-Passungsmaß

berechnen (AIST-R (.86 ≤ α ≤ .93) bzw. UST-R (.86 ≤ α ≤ .94); Bergmann & Eder, 2005). Zusätzlich wurde ein objektives Maß der kognitiven Grundfähigkeit erfasst (KFT; Heller & Perleth, 2000; α = .90).

Die Auswertungen basierten auf einem Nested-Faktor-Modell der direkten PUP, das einen Generalfaktor (positive Ladungen aller PUP-Items) sowie drei spezifische Faktoren („Interessen-Tätigkeiten-Passung", „Fähigkeiten-Anforderungen-Passung", „Person-Organisation-Passung") umfasste („Bedürfnis-Angebot-Passung" fungierte als Referenzfaktor; Eid et al., 2017). Dieses Modell wurde anschließend um die Kovariaten (indirekte Interessenkongruenz und kognitive Grundfähigkeit) erweitert. Es wurde überprüft, ob sich die Zusammenhänge der Kovariaten mit den direkten PUP-Maßen auf deren Zusammenhänge mit dem Generalfaktor reduzieren lassen, was dafür sprechen würde, dass die direkten Einschätzungen der PUP-Dimensionen eine globale Umwelteinschätzung widerspiegeln und ähnliche Assoziationen mit den Kovariaten aufweisen. Demgegenüber würden bedeutsame theoriekonforme Zusammenhänge der spezifischen Faktoren mit den Kovariaten („Interessen-Tätigkeiten-Passung" mit indirekter Interessenpassung; „Fähigkeiten-Anforderungen-Passung" mit kognitiver Grundfähigkeit) die Validität direkter PUP-Erhebungen stützen.

Ergebnisse und Diskussion:

Das Nested-Faktor-Modell wies eine gute Datenanpassung auf [$c^2(df = 45) = 319.8$; RMSEA = .054; CFI = .966; SRMR = .039]. Diese blieb nach Hinzunahme der Kovariaten weitestgehend unverändert. Die indirekt gemessene Interessenpassung korrelierte stark mit dem Generalfaktor ($r = .34$, $p < .001$) und wies keine nennenswerten Zusammenhänge mit den spezifischen Faktoren auf. Die kognitive Grundfähigkeit wies ebenso einen starken Zusammenhang mit dem Generalfaktor auf ($r = .18$, $p < .001$) und korrelierte außerdem mit den spezifischen Faktoren „Interessen-Tätigkeiten-Passung" ($r = .12$, $p < .001$) und „Fähigkeiten-Anforderungen-Passung" ($r = .22$, $p < .001$). Dieses Zusammenhangsmuster ist bestenfalls bedingt mit den theoretischen Erwartungen kompatibel.

Direkte PUP-Erfragungen (insbesondere die häufig verwendete „Interessen-Tätigkeiten-Passung") scheinen also die Gefahr zu bergen, nicht das intendierte Konstrukt zu messen.

Murat und Kevin – beide verhaltensauffällig und leistungsschwach? Inhaltliche Konstitution von Stereotypen über Schüler*innen mit türkischem Migrationshintergrund sowie Schüler*innen aus niedrigen sozialen Schichten bei (angehenden) Lehrkräften

Anna Shevchuk, Sabine Glock
Bergische Universität Wuppertal, Germany

Theoretischer Hintergrund und Fragestellung

Herkunftsbedingte Disparitäten im Bildungsverlauf sind immer noch aktuell (Maaz & Dumont, 2019). Die Suche nach den Ursachen zeigt, dass Schüler*innen mit einem (v. a. türkischen) Migrationshintergrund oder aus niedrigen sozialen Schichten von ihren Lehrkräften selbst unter Kontrolle der tatsächlichen Leistungen ungünstiger beurteilt und eingeschätzt werden als Schüler*innen ohne Migrationshintergrund sowie Schüler*innen aus höheren sozialen Schichten (Tobisch & Dresel, 2017). Gleichzeitig besteht ein starker Zusammenhang zwischen sozialer Herkunft und dem Migrationshintergrund.

Häufig werden Stereotype von Lehrkräften als ursächlich diskutiert. Stereotype umfassen Überzeugungen über Charakteristika, Eigenschaften und Verhaltensweisen von Mitgliedern bestimmter sozialer Gruppen (Hilton & von Hippel, 1996). Bislang ist bekannt, dass ein türkischer Migrationshintergrund mit einer häufigeren Zuschreibung negativer Stereotype (Froehlich et al., 2016) und ein niedriger sozialer Status allgemein mit negativen Attributen wie „faul“, „ungebildet“, „Migrationshintergrund“ (Lindqvist et al., 2017) assoziiert ist. Stereotype lassen sich dabei grundsätzlich auf zwei Dimensionen—Kompetenz und Wärme—einordnen (Fiske et al., 2002). So geht ein niedriger sozialer Status mit geringer Kompetenz, aber hoher Wärme einher (Durante et al., 2017) und Personen mit einem türkischen Migrationshintergrund werden eher mit sozialen als mit leistungsbezogenen Kompetenzen in Verbindung gebracht (Bonefeld & Karst, 2020). Gleichzeitig können Stereotype in ihrer Valenz variieren. Zu Personen mit einem türkischen Migrationshintergrund finden sich bei angehenden Lehrkräften insbesondere im Bereich der Leistung häufig negative Stereotype (Bonefeld & Karst, 2020); in anderen Bereichen hingegen rein neutrale Stereotype. Gegenüber Schüler*innen mit einem niedrigen sozialen Status finden sich tendenziell weniger positivere und mehr negativere Stereotype (Tobisch & Dresel, 2020), wobei die in der letztgenannten Studie untersuchten stereotypen Eigenschaften von den Forscher*innen selbst ausgewählt wurden und womöglich nicht die tatsächlich vorherrschenden Stereotype zu Schüler*innen aus unterschiedlichen sozialen Schichten abbilden. Vor diesem Hintergrund wird im Vortrag untersucht, ob sich die Stereotype von (angehenden) Lehrkräften über Schüler*innen mit einem türkischen Migrationshintergrund und Schüler*innen aus niedrigen sozialen Schichten unterscheiden.

Methode

Die Inhalte der Stereotype über Schüler*innen mit einem türkischen Migrationshintergrund und über Schüler*innen aus niedrigen sozialen Schichten wurden durch eine freie Assoziationsaufgabe erfasst. Die Probanden wurden gebeten, ihr Wissen über grundsätzlich unter den Lehrer*innen vorherrschenden Annahmen und Vorstellungen aufzulisten.

Ergebnisse

Die Stichprobe umfasste insgesamt 114 (angehende) Lehrkräfte. Es wurden sieben deduktive Kategorien („Materielles“, „Kultur und Gesellschaft“, „Charaktereigenschaften“, „Verhalten“, „Äußerliche Merkmale“, „Leistung“, „Persönlicher Bezug“), und vier induktive Kategorien („Sprache“, „Familie und Elternhaus“, „Lehrkraft (Verhalten)“ und „Reaktionen des Umfeldes“) verwendet. Alle Assoziationen wurden ebenso in die Kategorien positiv, negativ, neutral eingeordnet.

Über Schüler*innen mit einem türkischen Migrationshintergrund wurden mehr Assoziationen in den Kategorien „Kultur und Gesellschaft“, „Charakter“ und „Sprache“ genannt, wohingegen in den Kategorien „Materielles“, „Leistung“ und „Familie und Elternhaus“ mehr Nennungen über Schüler*innen aus niedrigen sozialen Schichten vorlagen. Zu Schüler*innen mit türkischem Migrationshintergrund wurden positive und neutrale, zu Schüler*innen aus niedrigen sozialen Schichten mehr negative Assoziationen genannt.

Diskussion

Die Stereotype über beide Schüler*innengruppen unterschieden sich in Anzahl und Inhalt. Die Assoziationen über Schüler*innen mit türkischem Migrationshintergrund und Schüler*innen aus niedrigen sozialen Schichten zeigten sich lediglich bei „Sprache“ (schlechte Deutschkenntnisse) und bei „Familie und Elternhaus“ als. überlappend, sind aber insgesamt hinreichend unterschiedlich.

Im Einklang mit der Forschung (Durante & Fiske, 2017) wurden bei Schüler*innen mit türkischem Migrationshintergrund positive Stereotype zum Charakter genannt. Kompetenzbezogen zeigte sich eine ähnliche Anzahl positiver und negativer Stereotype. Für Schüler*innen aus niedrigen sozialen Schichten fanden sich überwiegend Assoziationen negativer Valenz. Eine geringe Zuschreibung kompetenzbezogener Attribute wird jedoch nicht mit einer erhöhten Zuschreibung von Wärme kompensiert. Schüler*innen aus niedrigen sozialen Schichten verglichen mit Schüler*innen mit einem türkischen Migrationshintergrund könnten grundsätzlich mit mehr negativen Stereotypen von Lehrkräften konfrontiert sein. Diese negativen Stereotype können einen Einfluss auf die Urteile ausüben und folglich wesentlich zu der Erklärung herkunftsbedingter Disparitäten im Bildungsverlauf von Kindern aus unterschiedlichen sozialen Schichten beitragen.

Adaptive diagnostische Lehrkompetenz im inklusiven Unterricht – empirische Erfassung und theoretische Implikationen eines „dynamischen“ Konstrukts

Simone Dubiel, Lena Schmitz, Johanna Lau, Julia Frohn
Humboldt-Universität zu Berlin, Deutschland

Einleitung

Obwohl die Heterogenität von Lernenden in erziehungs- und bildungswissenschaftlicher Forschung seit jeher Relevanz hat, verlangen jüngste Entwicklungen – z. B. die Ratifizierung der UN-BRK 2009, die Corona-Pandemie, fluchtbedingte Sprachbarrieren – umso mehr nach einem inklusiven Unterricht, der den individuellen Voraussetzungen der Lernenden gerecht werden kann. In diesem Sinne wird inklusiver Unterricht verstanden als ein Recht aller Heranwachsenden auf qualitativ hochwertige Bildung, unabhängig des Geschlechts, der Religion, ethnischen Zugehörigkeit, der Lernbedürfnisse oder sozialen und ökonomischen Voraussetzungen (Salamanca Erklärung der UNESCO, 1994; UN-BRK, 2009). Gleichsam ist die Bestrebung eines gemeinsamen Unterrichts eine politisch gewollte und im Schulgesetz verankerte Verantwortung von Lehrkräften, um Chancengleichheit und Teilhabe aller Schüler:innen zu gewährleisten (vgl. z. B. Berliner SchulG, 2020).

Diese Verantwortung setzt jedoch voraus, dass Lehrkräfte den eigenen Unterricht in Planung und Durchführung bestmöglich auf die individuellen Bedürfnisse von Schüler:innen ausrichten können (Gold & Hasselhorn, 2017). Diese Fähigkeit kann als *adaptive Lehrkompetenz* bezeichnet (Beck, Baer, Guldimann, Bischoff & Brühwiler, 2008; Schmitz, 2017) und in Anlehnung an Helmke und Weinert (1997) in diagnostische, didaktische, Sach- sowie Klassenführungskompetenz untergliedert werden.

Forschungsdesiderat

Innerhalb des Projektes *Fachdidaktische Qualifizierung Inklusion angehender Lehrkräfte an der Humboldt-Universität zu Berlin - Schwerpunkt MINT-Bereich* (FDQI-HU-MINT) wurden Seminare zur

Förderung der adaptiven Lehrkompetenz von Lehramtsstudierenden entwickelt und empirisch evaluiert. Die Evaluation geht der Forschungsfrage nach:

Inwieweit verändert sich die adaptive diagnostische Lehrkompetenz von Studierenden innerhalb eines Semesters?

Methode

Die Stichprobe beruht auf Paneldaten von 84 Lehramtsstudierenden, die im Wintersemester 2021/22 im Prä-Post-Design mit Interventions- (*N*=79) und Vergleichsgruppe (*N*=5) erhoben wurden. Zur Erfassung der adaptiven Lehrkompetenz der Studierenden wurde ein Testinstrument (Schmitz, Simon & Pant, 2020), basierend auf Videovignetten und offenen Items, entwickelt. Anhand von Kodierschemata wurden Indizes berechnet, die jeweils einen Indexwert von 0 bis maximal 30 annehmen können. Die Ergebnisse wurden mithilfe einer Varianzanalyse mit Messwiederholung und unter Berücksichtigung der Kovariate Schulform (binäre Variable Grundschule/weiterführende Schule) (ANCOVA) mit dem Statistikprogramm SPSS (IBM Corp., 2022) ausgewertet.

Ergebnisse und Diskussion

Die Ergebnisse zeigen, dass der durchschnittliche Indexwert der adaptiven diagnostischen Lehrkompetenz zum ersten Messzeitpunkt für die Interventionsgruppe bei 9.22 (SD = 4.05) liegt und sich zum zweiten Messzeitpunkt um 2.29 Indexpunkte auf 6.92 (SD = 3.61) verringert. Die Daten der Vergleichsgruppe zeigen ein ähnliches Muster. Nach Bereinigung um das angegebene Lehramt (Grundschullehramt bzw. Lehramt einer weiterführenden Schule) unterscheidet sich der Index der adaptiven diagnostischen Lehrkompetenz nicht statistisch signifikant für die zwei Messzeitpunkte: $F(1, 80) = 2.603$, $p > .05$, partielles $\eta^2 = .032$. Vergleichsdaten zur adaptiven didaktischen Lehrkompetenz und Klassenführungskompetenz werden aktuell analysiert und miteinander kontrastiert.

Insgesamt gilt das Konstrukt der adaptive Lehrkompetenz als „komplex“ und „dynamisch“ (Prengel, 2016; Schmitz, Brodesser & Pant, 2020). Darüber hinaus können methodische Herausforderungen von Panelstudien (z. B. „panel conditioning“, siehe Warren & Halpern-Manners, 2012) diskutiert werden. Die Erfassung des Konstrukts ist gleichzeitig herausfordernd und notwendig, um auf Grundlage empirischer Daten evidenzbasiert diskutieren zu können. Diese Diskussion soll im Vortrag um eine theoretische Komponente ergänzt werden: So hat die wiederholte Messung der adaptiven Lehrkompetenz im Projekt FDQI-HU auch zu einer theoretischen (Neu-)Verortung des Konstrukts zwischen kompetenzorientierten und strukturtheoretischen Ansätzen der Lehrkräfteprofessionalisierung geführt. Diese (Neu-)Verortung wird als Ausblick vorgestellt und eröffnet weiterführende Fragen zur empirischen Erfassung und theoretischen Fundierung des Konstrukts.

F-08: Einstellungen und Überzeugungen von Lehrkräften

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S05 T00 B71
Chair der Sitzung: Tobias Hoppe

Einfluss von Wissen und eigenen Einstellungen auf das Diagnostizieren von Lernendeneinstellungen zu Evolution und Schöpfung durch angehende Biologie- und Religionslehrkräfte

Tobias Hoppe[1], Christiane Konnemann[2], Christian Höger[3], Alexander Renkl[4], Werner Rieß[1]
[1]Pädagogische Hochschule Freiburg, Deutschland; [2]Westfälische Wilhelms-Universität, Münster, Deutschland; [3]Universität Luzern, Schweiz; [4]Albert-Ludwigs-Universität, Freiburg, Deutschland

Theoretischer Hintergrund

Einstellungen und Überzeugungen von Lernenden in Bezug auf das Verhältnis von Schöpfungsglaube und Evolutionstheorie sind divers und werden aus fachdidaktischer sowie wissenschaftstheoretischer Perspektive teilweise als problematisch betrachtet (Höger, 2020; Konnemann et al., 2016; Rothgangel, 2017; 2019). Das Diagnostizieren von Einstellungen in diesem Kontext stellt für Lehrkräfte eine besondere Herausforderung dar, weil beim Diagnostizieren auf Wissen aus unterschiedlichen Domänen zurückgegriffen werden muss (z. B. *Nature of Science*). Darüber hinaus haben sich moderierende Wirkungen von Personencharakteristika der urteilenden Lehrkräfte auf die Diagnosen gezeigt: So können eigene Einstellungen und Überzeugungen der Lehrkräfte (Klose, 2014), aber auch fachwissenschaftliches und fachdidaktisches Wissen (van Dijk & Kattmann, 2010) einen Einfluss auf die Urteile von Lehrkräften nehmen.

Als Querschnittsthema betrifft das Verhältnis von Evolution und Schöpfung sowohl den Biologie- als auch den Religionsunterricht (Gemballa & Schweitzer, 2010). Es ist davon auszugehen, dass sich angehende Biologie- und Religionslehrkräfte u.a. aufgrund unterschiedlicher Wissensbestände in ihren Diagnosefertigkeiten unterscheiden. Insbesondere in der Ausbildung von Biologielehrer*innen spielen wissenschaftsphilosophische Reflexionen kaum eine Rolle (Dittmer & Zabel, 2019). Um diagnostische Fertigkeiten bei den (angehenden) Lehrkräften gezielt fördern zu können, gilt es ein Verständnis davon zu entwickeln, wie Lehrkräfte zu diagnostischen Urteilen kommen (Loibl, Leuders, & Dörfler, 2020).

Fragestellung und Methode

In der laufenden Studie wird untersucht, wie angehende Biologie- und Religionslehrer*innen vor dem Hintergrund eigener Personencharakteristika (u. a. Wissen, eigene Einstellungen und Überzeugungen, Fachlehrer*innensein) zu diagnostischen Urteilen über Lernendeneinstellungen zu Schöpfungsglaube und Evolutionstheorie kommen. Die diagnostischen Urteilsprozesse des Wahrnehmens und Interpretierens werden über einen Vignettentest erfasst. Den Versuchspersonen werden textbasierte Unterrichtssequenzen gezeigt, in denen bestimmte Äußerungen von Schüler*innen auf deren Einstellungen zu Evolutionstheorie und Schöpfungsglaube schließen lassen. Diese Äußerungen wurden als Hinweisreize in den Vignetten systematisch variiert (Schnitzler, Holzberger, & Seidel, 2020). Die inhaltliche Validität wurde mithilfe eines Expertenratings (N=10) überprüft. Es konnte eine sehr gute Urteilerübereinstimmung festgestellt werden (ICC=0,983). Ferner wurde die Authentizität der abgebildeten Unterrichtsauschnitte durch Lehrkräfte im Schuldienst (N=11) eingeschätzt und im Mittel als zufriedenstellend bezeichnet. Im Rahmen des Vignettentests ermöglicht ein offenes Aufgabenformat die Bestimmung unterschiedlicher Niveaus diagnostischer

Fertigkeiten (Hoppe, Renkl, & Rieß, 2020). Die Beobachtungen der Proband*innen werden in Anlehnung an die qualitative Inhaltsanalyse nach Mayring ausgewertet.

Diagnoserelevante Personencharakteristika werden mithilfe eines Fragebogens über offene und geschlossene Items erfasst. Dabei wird weitgehend auf Instrumente mit belegter Reliabilität und Validität zurückgegriffen (z. B. CAEB, Stahl & Bromme, 2007; Kurzskalen zur Erfassung von szientistischen und kreationistischen Überzeugungen, Konnemann et al., 2016). Ausgehend von Korrelationsanalysen sollen Schlüsse dahingehend gezogen werden, wie sich bestimmte Personencharakterisika der diagnostizierenden Lehrkräfte auf die Urteilsbildung auswirken. An der Untersuchung haben 110 Lehramtsstudierende teilgenommen (n=63 mit Unterrichtsfach Biologie; n=47 mit Unterrichtsfach Ev. oder Kath. Religionslehre/Religionspädagogik).

Ergebnisse und ihre Bedeutung

Hinsichtlich der Interpretation wahrgenommener Äußerungen zeigte sich, dass Lehramtsstudierende zentrale Hinweisreize bei den Lernenden zwar bei der Urteilsbildung verarbeiteten, aber teilweise nicht angemessen deuten konnten. Mithilfe von t-Tests konnten dabei signifikante Unterschiede in der Interpretationsleistung von angehenden Biologie- und Religionslehrkräften z.B. beim Diagnostizieren von kreationistischen Einstellungsprofilen nachgewiesen werden, $t(103) = -2.91$, $p = .004$. Angehende Religionslehrkräfte interpretierten die entsprechenden Einstellungsprofile insgesamt zutreffender. Korrelationsanalysen weisen darauf hin, dass Lehrkräfte die Einstellungen von Lernenden tendenziell dann besonders zutreffend interpretieren können, wenn ihre eigene Wahrnehmung eines Konflikts zwischen Naturwissenschaften und Religion niedrig ist. Die Interpretationsleistung angehender Lehrkräfte bezogen auf kreationistische Einstellungsprofile und eigene Überzeugungen zur Wahrnehmung eines Konflikts zwischen Naturwissenschaften und Theologie korrelierte moderat negativ, $r = -.289$, $p < .01$. Eine weiterführende Analyse der Daten wird in den nächsten Wochen erfolgen und im Vortrag präsentiert werden. Die Ergebnisse der Untersuchung dienen dazu, eine Intervention zur Förderung der Diagnosefertigkeiten von angehenden Lehrkräften in den Fächern Biologie und Religionslehre zu entwickeln.

Die Bedeutsamkeit des Fachinteresses für die Selbstwirksamkeitserwartungen von Lehrer:innen im Bereich der politischen Bildung

Hermann Josef Abs, Ines Leonie Laforet
Universität Duisburg-Essen, Germany

Bereits der Kritiker Walter Pater (1839-1894) erkannte die Bedeutung des Wissens um die eigenen Interessen als grundlegende Voraussetzung dafür, dieses Interesse auch bei anderen zu wecken bzw. sie damit anzustecken. Der Vortrag stellt das Fachinteresse an Politik und Wirtschaft von Lehrkräften der Sozialwissenschaften in den Mittelpunkt und untersucht dessen Auswirkungen auf Selbstwirksamkeitserwartung. Bisherige Studien haben im Hinblick auf das Fachinteresse vor allem die allgemeine Selbstwirksamkeitserwartung bzw. die Selbstwirksamkeitserwartungen in den Bereichen der Lernförderung oder Unterrichtsgestaltung betrachtet (vgl. Ekstam et al., 2017; Oberle/ Weschenfelder/ Weißeno, 2013). In der vorliegenden Untersuchung wird dagegen die erwartete Selbstwirksamkeit der Lehrer:innen bezogen auf das Unterrichten verschiedener Themen und Fähigkeiten im Bereich der politischen Bildung gemessen. Diese zeichnet sich im Vergleich zu

anderen Fächern durch ein weit gefasstes Themenfeld aus, wodurch Unsicherheiten bei den Lehrer:innen entstehen und deren eigene Erwartung der erfolgreichen Lehrbewältigung gemindert werden können (vgl. Schwarzer/ Warner, 2014).

Die Untersuchung erweitert die bisherige Forschung durch eine differenziertere Konzeptualisierung und Erfassung des Lehrer:inneninteresses. Bislang untergliederten Studien das Interesse der Lehrer:innen in ein Fach-, ein fachdidaktisches und ein pädagogisches Interesse (vgl. Pozas/ Letzel, 2021; Schiefele/ Streblow/ Retelsdorf, 2013). Fachdidaktische Studien konzentrierten sich im Zusammenhang mit dem Interesse vor allem auf das Fachinteresse Politik (vgl. Weißeno/ Weschenfelder/ Oberle, 2013). In der vorliegenden Arbeit wird dagegen das Fachinteresse zur politischen Bildung weiter ausdifferenziert. Neben dem Einfluss des Fachinteresses Politik wird untersucht, welche Wirkung das Fachinteresse Ökonomie auf die Selbstwirksamkeit hat, da davon ausgegangen wird, dass sich die Wirkungsweise der beiden Fachinteressen unterscheidet. Die gewählte Differenzierung der beiden fachlichen Interessen lässt nämlich folgende Beschreibung zu: Ein hohes politisches Interesse ist mit einer progressiven Sicht auf die Welt und einem hohen Maß an Altruismus assoziiert, während ein hohes ökonomisches Interesse mit einer stärker konservativen Sichtweise und der Verfolgung von Eigeninteressen verbunden ist (vgl. Fischer et al., 2016; Gandal et al., 2005).

Die Datengrundlage für die Untersuchung bilden die Daten aus dem Feldtest der „International Civic and Citizenship Education Study 2022“. Um die Bedeutung des Fachinteresses für die Selbstwirksamkeitserwartungen der Lehrer:innen im Bereich der politischen Bildung zu messen, wird eine hierarchische multiple Regressionenanalyse eingesetzt. Da zahlreiche Studien bezogen auf die Selbstwirksamkeitserwartung einen Zusammenhang mit dem Alter und dem Geschlecht sowie mit der Aus- bzw. Fortbildungsintensität feststellen (vgl. Pozas/ Letzel, 2021; Weißeno/ Weschenfelder/ Oberle, 2014; Wudy/ Jerusalem, 2011), bilden diese drei Faktoren die unabhängigen Kontrollvariablen. Um die Variablen der Selbstwirksamkeitserwartungen zu konstruieren wurde vorbereitend eine Faktorenanalyse durchgeführt, die zu drei Faktoren führte. Die Selbstwirksamkeit wird entsprechend in den Dimensionen (politische Institutionen, politischer Prozess und Ökonomie) untergliedert.

Im Ergebnis zeigt die Untersuchung überwiegend positive Zusammenhänge zwischen Fachinteressen und Selbstwirksamkeitserwartungen. In den Dimensionen politische Institutionen und politischer Prozess wirken beide Fachinteressen positiv. Da die beiden Konstrukte Fachinteresse und Selbstwirksamkeit eine gewisse Nähe zueinander aufweisen, da sie zwei Facetten der motivationalen Merkmale von Lehrkräften sind, war dies durchaus zu erwarten. Die Ergebnisse in der Regression zur Dimension Ökonomie deuten jedoch eine Interessensdifferenzierung an. So wirkt das ökonomische Interesse (.690**) positiv auf die Selbstwirksamkeitserwartungen im ökonomischen Bereich, während das politische Interesse diesbezüglich einen negativen Einfluss aufweist (-.174*).

Das Unterrichten von wirtschaftlichen Theorien bewegt sich im Spannungsfeld zwischen deontologischen und individuellen Ansprüchen. Lehrer:innen mit hohem politischen Interesse besitzen dabei möglicherweise einen inneren Konflikt durch ihre vorherrschenden altruistischen Werte. Folglich

ist es essenziell, zukünftige Lehrkräfte in der Lehrer:innenbildung für beide Bezugsdisziplinen gleichermaßen zu interessieren. Einer solipsistischen Betrachtung der Disziplinen sollte entgegengewirkt werden, sodass Lehrkräfte die Bereiche nicht als unabhängig voneinander, sondern vielmehr als interdependent wahrnehmen und anerkennen. Denn die aktuelle politische Lage beweist erneut die Notwendigkeit einer integrierten Betrachtung von Politik und Ökonomie.

Überzeugungen von Lehrkräften sozialwissenschaftlichen Fachunterrichts zur Gestaltung von Kontroversität im Unterricht

Katrin Hahn-Laudenberg[1], Hermann-Josef Abs[2], Johanna Fee Ziemes[2]
[1]Universität Leipzig, Germany; [2]Universität Duisburg-Essen

Theoretischer Hintergrund

Kontroversität ist als zentrales Qualitätsmerkmal insbesondere sozialwissenschaftlichen Unterrichts unbestritten (Abs, 2005; Grammes, 2020). Die Polarisierung öffentlicher Diskurse und die damit zum Teil verbundene Infragestellung von Rationalitätskriterien, aber auch von demokratischen Grundwerten fordern Lehrkräfte bei der Gestaltung von Kontroversität im Unterricht verstärkt heraus. Die gestiegene Relevanz spiegelt sich in einem intensivierten konzeptuellen Diskurs über Kriterien von Kontroversität (Drerup et al., 2021). Im deutschsprachigen Raum gibt es bisher allerdings kaum qualitative (Behrens, 2014) und im noch geringere Maße quantitativ gestützte Erkenntnisse (Oberle et al., 2018, Autor:in, eingereicht) zum Umgang zu Kontroversität im Unterricht und darauf bezogene Überzeugungen von (angehenden) Lehrkräften. Diese beziehen sich zudem bislang vor allem auf die Frage eines vermeintlichen Neutralitätsgebots sowie den Umgang mit extremistischen Positionen und damit eher auf Grenzen des Kontroversitätsgebots als auf dessen Realisierung. Letztere ist stärker im Fokus eines internationalen Forschungsdiskurses. Mit Hess (2009) lassen sich bei Lehrkräften vier Typen des Umgangs mit Themen, die wissenschaftlich und politisch kontrovers sind, unterscheiden: Das Vermeiden kontroverser Themen (Avoidance), Negieren/Leugnen des kontroversen Gehalts eines Themas (Denial) oder die Privilegierung einer Position (Privilege) können von einer grundsätzlich anzustrebenden ausgewogenen Darstellung kontroverser Positionen (Balance) unterschieden werden. Ihre, sowie daran anschließende, internationale ethnographische Forschung geben empirische Hinweise auf die Virulenz von Vermeidungsverhalten bei Lehrkräften, das sich unter anderem in einem Ausweichen auf übergeordnete (etwa werteverbinde) Grundlagen äußern kann (Pollak et al., 2018) und von Lehrkräften selbst etwa mit einer Angst vor Konflikten und Kontrollverlust sowie dem Gefühl der Überforderung angesichts der Komplexität kontroverser Themen begründet wird (Flensner, 2020). Hierauf bezogene Konzepte der Lehrkraftbildung basieren auf der Annahme, dass Lehrkraftausbildung einen professionellen Umgang mit Kontroversität fördern kann (Pace, 2021).

Fragestellung

Welche Überzeugungen zur Gestaltung von Kontroversität im Unterricht äußern Lehrkräfte, die in sozialwissenschaftlichen Fächern unterrichten? Welche systematischen Unterschiede zwischen unterschiedlichen Lehrkräften sind erkennbar?

Methode

Die Analysen basieren auf Daten der nationalen Ergänzungsstichprobe der Lehrkraftbefragung im Rahmen der *International Civic and Citizenship Education Study* 2022 der IEA. Die Onlinebefragung fand in der ersten Jahreshälfte 2022 an 146 Schulen in Nordrhein-Westfalen und 92 Schulen in Schleswig-Holstein statt. In der nationalen Ergänzungsstichprobe wurden Lehrkräfte befragt, die in sozialwissenschaftlichen Fächern eingesetzt werden, aber nicht von der internationalen Stichprobe abgedeckt waren. Die ausgewertete Stichprobe umfasst N=831 Lehrkräfte. Analysiert wird inwiefern ein neu entwickeltes Instrument in der Lage ist, die Kontroversitätstypen Avoidance, Privilege und Balance nach Hess (2009) zu operationalisieren (CFA in Mplus 8.2). In einem zweiten Schritt wird mittels Regressionsrechnungen die Bedeutung von individuellen Merkmalen (u.a. staatliche Lehrbefähigung) und schulspezifischen Faktoren (u.a. Schulform) geprüft.

Ergebnisse und ihre Bedeutung

Die CFA bestätigt die dreidimensionale Struktur, welche mit Blick auf die Faktorladungen jedoch nur hinsichtlich der Vermeidung von Kontroversität im Unterricht (Avoidance) die Bildung einer Skala (6 Items, α=.781) nahelegt. Die Regressionsanalyse (N=599) zeigt, dass eine fehlende staatliche Lehrbefähigung (*B*=.135**), sowie (im Vergleich zum Gymnasium) die Arbeit an einer Gesamtschule (*B*=.137*), und insbesondere an einer nicht zum Abitur führenden Schulform (*B*=.294***) Vermeidungsverhalten wahrscheinlicher werden lässt. Die Ergebnisse bleiben auch bei Berücksichtigung anderer Schulmerkmale (Größe, Anteil MGH) stabil.

Die Bedeutung schulformspezifischer Kontexte sowie der Lehrbefähigung für eine ausgewogene Gestaltung kontroverser Themen (Balance) bestätigen zusätzliche Analysen von einem Item zur aktiven Herstellung von Kontroversität durch die Lehrkraft in politisch einseitig positionierten, homogenen Lerngruppen.

Die Ergebnisse liefern empirische Evidenz für die Relevanz einer sozialwissenschaftsspezifischen Lehrerbildung und insbesondere für die Ausbildung qualitätsrelevanter Überzeugungen zur Gestaltung von Kontroversität. Entsprechend wird die Befürchtung gestützt, dass der breite Einsatz fachfremder Lehrkräfte in sozialwissenschaftlichen Fächern Einbußen in zentralen Aspekten sozialwissenschaftsdidaktischer Unterrichtsqualität mit sich bringt. Die darüberhinausgehende Bedeutung der Schulformen weist auf differentielle Lernumgebungen hin, welche Schüler*innen an Haupt- Real- Sekundar- und Förderschulen mit Blick auf die Möglichkeiten der Entwicklung politischer Kompetenzen benachteiligen.

F-09: Wohlbefinden von Lehrkräften

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S05 T00 B83
Chair der Sitzung: Manuela Haldimann

Das Wohlbefinden von Lehramtsstudierenden in Deutschland, Österreich und der Schweiz – Eine latente Profilanalyse

Manuela Haldimann[1], Tina Hascher[2], Doreen Flick-Holtsch[3]
[1]Pädagogische Hochschule St.Gallen, Schweiz; [2]Universität Bern, Schweiz; [3]Universität Zürich, Schweiz

Theoretischer Hintergrund

Der hohe Lehrpersonenbedarf in Kombination mit einer hohen Berufsausstiegsrate – insbesondere bei Junglehrpersonen – ist in Bildungssystemen weltweit ein Thema (OECD, 2020). In diesem Kontext gewinnt das Wohlbefinden von Lehrpersonen nicht nur in der Forschung (Hascher & Waber, 2021; McCallum et al., 2017), sondern auch in der Bildungspolitik (Viac & Fraser, 2020) zunehmend an Wichtigkeit. Studien, in denen das Wohlbefinden angehender Lehrpersonen empirisch untersucht wird, sind bisher rar (Bjorklund et al., 2021; Corcoran & O'Flaherty, 2022). Dies ist erstaunlich vor dem Hintergrund, dass das Wohlbefinden einen zentralen Teil des Professionsverständnisses ausmacht und die Grundsteine für eine kontinuierliche Professionalisierung bereits während der Ausbildung gelegt werden (Herzog et al., 2021). Ziel der Studie ist deshalb, das Wohlbefinden von Lehramtsstudierenden mehrdimensional, kontextspezifisch sowie personenzentriert zu untersuchen. Abgestützt auf das theoretische Verständnis der Wohlbefindenspsychologie (Becker, 1994; Diener, 1984) wird das Wohlbefinden der Lehramtsstudierenden als längerfristige Dominanz von positiven Emotionen, Kognitionen und physischen Empfindungen gegenüber negativen Emotionen, Kognitionen und physischen Empfindungen in Bezug auf das Lehramtsstudium definiert. Basierend auf dem Modell schulischen Wohlbefindens von Hascher (2004) wird das Wohlbefinden von Lehramtsstudierenden mit drei positiven Dimensionen (1) Positive Einstellungen zum Studium, (2) Freude im Studium, (3) Akademischer Selbstwert, sowie drei negativen Dimensionen (4) Sorgen wegen des Studiums, (5) Physische Probleme wegen des Studiums und (6) Soziale Probleme im Studium konzeptualisiert. Bisherige Forschungsbefunde lassen Zusammenhänge des Wohlbefindens von angehenden Lehrpersonen mit der Studienabbruchintention (Klassen & Chiu, 2011) und ihrer Resilienz (Hascher et al., 2021), später mit ihrer Berufszufriedenheit (Capone & Petrillo, 2018) vermuten.

Fragestellungen

Im Rahmen des Beitrags wird drei Fragestellungen nachgegangen:

1. Wie ist das Wohlbefinden von Lehramtsstudierenden auf den sechs Wohlbefindensdimensionen ausgeprägt?
2. Welche Wohlbefindensprofile lassen sich bei Lehramtsstudierenden identifizieren?
3. Inwiefern unterscheiden sich die identifizierten Profile in Bezug auf die Studienabbruchintention, die Resilienz sowie die antizipierte Berufszufriedenheit?

Methode

Datengrundlage bilden die Antworten von 1749 Lehramtsstudierenden aus Deutschland, Österreich und der Schweiz, welche im Frühjahr 2022 an der Onlinebefragung «Lehrperson werden: Wohlbefinden im Studium» teilgenommen haben. In der Onlinebefragung wurden neben dem Wohlbefinden (adaptiert von Hascher, 2020) die Studienabbruchintention (Kunter et al., 2017), die Resilienz (Schumacher et al., 2005) sowie die antizipierte Berufszufriedenheit (Baumert et al., 2009) je mit einer sechsstufigen Likert-Skala erfasst. Zur Beantwortung der ersten Fragestellung wurden die

Daten deskriptiv analysiert. Anschliessend wurde in Mplus (Version 8.7) eine latente Profilanalyse durchgeführt und untersucht, inwiefern sich zwischen den dabei identifizierten Profilen signifikante Unterschiede bezüglich der Resilienz, der Studienabbruchintention sowie der antizipierten Berufszufriedenheit zeigen. Entsprechend dem «Three-Step Approach» nach Asparouhov and Muthén (2014) wurde zur Berechnung der Unterschiede nicht von einer eindeutigen Profilzugehörigkeit ausgegangen, sondern die Unsicherheit der Klassifizierung mitberücksichtigt.

Ergebnisse und ihre Bedeutung

Erste Analysen zeigen, dass die Reliabilität, Validität sowie die (partielle) Messinvarianz des Erhebungsinstruments zur Erfassung des Wohlbefindens von Lehramtsstudierenden gegeben ist, wobei der Faktor «Soziale Probleme im Studium» aufgrund einer niedrigen Reliabilität ausgeschlossen wird (α = .50). Bei den positiven Dimensionen resultieren für das Wohlbefinden tendenziell günstige Werte (*M*min= 3.45; *M*max = 4.41; *SD*min = 0.79; *SD*max = 1.00). Während dies ebenfalls auf die negative Dimension «Physische Probleme wegen dem Studium» zutrifft (*M* = 2.64; *SD* = 1.35), zeigen sich eher ungünstige, also hohe, Werte bezüglich «Sorgen wegen des Studiums» (*M* = 4.04; *SD* = 1.25). Es konnten vier verschiedene Profile identifiziert werden, welche die Basis für die weiterführenden Analysen bilden. Neben einem Profil mit hohem Wohlbefinden (34.5 %) und einem Profil mit niedrigem Wohlbefinden (12.0 %) lassen sich auch zwei mittlere Profile (44.3 % bzw. 9.3 %) mit je spezifischen Ausprägungskombinationen charakterisieren.

Die Ergebnisse werden im Hinblick auf eine mehrdimensionale Erfassung des Wohlbefindens angehender Lehrpersonen verschiedener Ausbildungssysteme diskutiert und sollen Impulse für die Wohlbefindensförderung in der Ausbildung von Lehrpersonen geben.

individuelle und äußere Einflussfaktoren auf das Wohlbefinden angehender Lehrkräfte

Stefanie Gäckle, Claudia Menge, Andreas Ortenburger
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), Deutschland

Ausgehend davon, dass Belastung und Zufriedenheit im Lehrkraftberuf sowohl von den Arbeitsbedingungen in Schule und Referendariat abhängen als auch davon, wie die (angehenden) Lehrkräfte ihre Situation individuell wahrnehmen und darauf reagieren, werden in diesem Beitrag verschiedene individuelle und berufliche Faktoren hinsichtlich ihres Einflusses auf das Wohlbefinden von angehenden Lehrkräften untersucht.

Das Referendariat gilt als wichtigste formale Lernumwelt zur Entwicklung professioneller Kompetenzen von Lehrkräften (Kunter et al. 2011). Inwiefern die Bedingungen im Referendariat, beispielsweise die Unterstützung durch MentorInnen, Seminarleitung und MitreferendarInnen, das Wohlbefinden von (angehenden) Lehrkräften beeinflussen, soll der erste Teil dieses vorgeschlagenen Beitrags sein. Trotz intensiver Forschung zur beruflichen Belastung von Lehrkräften liegen kaum belastbare Studien zum Einfluss der Lernumwelt auf die emotionale Befindlichkeit von (angehenden) Lehrkräften vor. Die wenigen vorliegenden Studien zum Thema beschränken sich auf einzelne Bundesländer (z. B. Nübling et al. 2012; Drüge et al. 2014) oder unterrichtete Fächer (z. B. Klusmann et al. 2008b).

Neben diesen Lernumweltsmerkmalen wird eine Skala zur beruflichen Selbstregulation berücksichtigt, die das individuelle Erleben und Verhalten in Bezug auf berufliche Belastungen widerspiegelt. Berufliche Selbstregulation ist definiert als die Fähigkeit, mit den eigenen Ressourcen bei der Arbeit effektiv zu haushalten (Kunter et al. 2011), um ein dauerhaftes berufliches Engagement aufrechtzuerhalten und sich in der Freizeit zu erholen. Die eingesetzten Items basieren auf dem AVEM-Inventar von Schaarschmidt und Fischer (2001). Die auf Basis des Inventars ermittelbaren AVE-Muster gelten als relativ stabile Dispositionen und sollen der frühzeitigen Identifikation von Warnsignalen dienen, die bei auftretenden Belastungssituationen zu gesundheitlichen Beeinträchtigungen führen können (vgl. Schaarschmidt und Kieschke 2007). Sie stellen relevante Prädiktoren für die Gesundheit und Unterrichtsqualität von Lehrkräften dar (Schaarschmidt & Fischer 2001).

Die Analysen basieren auf Daten der Startkohorte 5 des Nationalen Bildungspanels, die seit dem Wintersemester 2010/2011 eine bundesweite Zufallsstichprobe von Studienanfängerinnen und -anfängern an deutschen Hochschulen durch das Studium und bis in das Berufsleben begleitet (NEPS Netzwerk 2022). Die Analysestichprobe umfasst (angehende) Lehrkräfte, die den ersten Teil der Lehrkräfteausbildung abgeschlossen haben und über gültige Daten zu allen relevanten Konstrukten verfügen (N=775).

Die Lernumwelt Referendariat wird erfasst anhand der instrumentellen Unterstützung durch MentorInnen, SeminarleiterInnen und MitreferendarInnen (jeweils drei Items; $\alpha_m = 0{,}79$; $\alpha_s = 0{,}72$; $\alpha_p = 0{,}80$; Kunter et al. 2016) sowie der Interaktion mit MentorInnen (konstruktivistischer Interaktionstyp; vier Items, $\alpha = 0{,}78$; transmissiver Interaktionstyp; drei Items, $\alpha = 0{,}73$; Max-Planck-Institut für Bildungsforschung 2010). Zur Messung der beruflichen Selbstregulation wurde eine 13 Items umfassende Kurzskala der arbeitsbezogenen Verhaltens- und Erlebensmuster (AVEM; Schaarschmidt & Fischer 2001) eingesetzt. Diese umfasst jeweils zwei Dimensionen zum Arbeitsengagement und der Widerstandsfähigkeit gegenüber beruflichen Belastungen. Das Wohlbefinden wird anhand der emotionalen Erschöpfung und der Arbeitszufriedenheit gemessen. Die Skala zur emotionalen Erschöpfung enthält vier Items aus einer deutschen Übersetzung des Maslach Burnout Inventory (Maslach et al. 1986; $\alpha = 0{,}80$). Die Arbeitszufriedenheit wurde anhand von vier Items von Kunter et al. (2011; $\alpha = 0{,}84$) erfasst. Für alle Befragten wird jeweils die erste Messung nach der Aufnahme des Referendariats berücksichtigt.

In Voranalysen wurden die Faktorstruktur der latenten Konstrukte über konfirmatorische Faktorenanalysen überprüft und die Muster beruflicher Selbstregulation über eine latente Profilanalyse generiert. Anschließend wurde der Einfluss der Lernumwelt sowie der beruflichen Selbstregulation auf das Wohlbefinden angehender Lehrkräfte mittels Strukturgleichungsmodellierung (SEM) analysiert. Das Modell weist einen zufriedenstellenden Modellfit auf (RMSEA=0,036, CFI=0,957, TLI=0,949, SRMR=0,038) und zeigt signifikante Effekte der Interaktion mit MentorInnen und Selbstregulationsfähigkeit. In der Präsentation wird das Modell um ausgewählte Kontrollvariablen wie die Schulform und das Geschlecht ergänzt. Darüber hinaus wird in einem Gruppenvergleich überprüft, ob sich die Interaktion mit und Unterstützung durch MentorInnen, Seminarleitung und

MitreferendarInnen zwischen den einzelnen Selbstregulationstypen unterschiedlich auf das Wohlbefinden auswirkt. Die Ergebnisse werden vorgestellt und diskutiert.

Wohlbefinden von Lehrkräften im Vorbereitungsdienst – die Rolle individueller Ressourcen

Elena Hohensee[1], Kira Elena Weber[2], Stephan Schiemann[1,3]
[1]Zukunftszentrum Lehrkräftebildung: ZZL-Netzwerk, Leuphana Universität Lüneburg, Germany; [2]Erziehungswissenschaft und Pädagogische Psychologie, IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany; [3]Institut für Bewegung, Sport und Gesundheit, Leuphana Universität Lüneburg, Germany

Einleitung

Das Wohlbefinden von (angehenden) Lehrkräften hängt eng mit ihren individuellen Ressourcen zusammen. Eine im Rahmen der Lehrkräftegesundheit noch wenig beforschte Ressource stellt die *individuelle Gesundheitskompetenz* (Soellner et al., 2017) dar, die auf der Ebene von Schlüsselkompetenzen (Weinert, 2001) die folgenden inhaltlichen Komponenten umfasst: *Selbstregulation, Selbstkontrolle, Selbstwahrnehmung, Verantwortungsübernahme, Umgang mit Gesundheitsinformationen, Kommunikation und Kooperation*. Sie wird als zu stärkende Ressource verstanden, die für die Gesunderhaltung und –förderung essentiell ist (Nutbeam, 1998). Erste Studien verweisen auf eine eingeschränkte Gesundheitskompetenz bei Lehramtsstudierenden und Lehrkräften (z.B. Hartmann et al., 2020; Ahmadi & Montazeri, 2019). Empirische Befunde zu Lehrkräften im Vorbereitungsdienst (LiV) liegen bislang noch nicht vor. Eine individuelle Ressource, die bislang nur im Kontext der beruflichen Gesundheit von (angehenden) Lehrkräften beforscht wurde (z. B. Klusmann et al., 2008), ist die *berufliche Selbstregulation* (Baumert & Kunter, 2011), die von Rupprecht (2015) auch als 'professionelle Gesundheitskompetenz' beschrieben wird. Aktuelle Studienergebnisse zeigen, dass bereits ca. 40% der LiVs gesundheitsgefährdende Bewältigungsmuster aufweisen (z. B. Lohse-Bossenz & Rutsch, 2021). Aus einer *Public Health Perspektive* sollten die beiden individuellen Ressourcen nicht getrennt voneinander betrachtet werden, sondern können gemeinsam einen Beitrag zur Gesundheitsförderung von angehenden Lehrkräften leisten.

Fragestellungen

Daran anknüpfend untersucht diese Studie folgende Fragestellungen:

1. Wie ist die *individuelle Gesundheitskompetenz* bei LiVs ausgeprägt?
2. Wie ist die *berufliche Selbstregulation* in Form von Bewältigungsmustern (Selbstregulationstypen) bei LiVs ausgeprägt?
3. Wie hängen die beiden individuellen Ressourcen zusammen und welchen Beitrag leisten sie zur Aufklärung des subjektiven Wohlbefindens von LiVs.

Methode

407 LiVs (weiblich=82.6%) nahmen von November 2020 bis Februar 2021 an der anonymen Online-Befragung teil. Die *individuelle Gesundheitskompetenz* wurde auf der Grundlage des Fragebogens von Lenartz (2011) gemessen, das Antwortformat hierbei ist vierstufig (1=*trifft überhaupt nicht zu* bis 4=*trifft genau zu*). Die *berufliche Selbstregulation* wurde mittels des AVEM-Fragebogens (Schaarschmidt & Fischer, 2008) auf einer fünfstufigen Likert-Skala (1=*trifft überhaupt nicht zu* bis 5=*trifft völlig zu*) erfasst und das *subjektive Wohlbefinden* mit dem WHO-5 Fragebogen auf einer sechsstufigen Likert-Skala (0=*zu keinem Zeitpunkt* bis 5=*die ganze Zeit*) (Krieger et al., 2014). Zur Beantwortung der Forschungsfragen wurden zum einen deskriptive Kennwerte herangezogen und mit

Verweis auf Klusmann et al. (2008) die arbeitsbezogenen Bewältigungsmuster mittels einer latenten Profilanalyse repliziert. Des Weiteren wurden Korrelationen, ANOVAs mit Post-hoc-Tests und multiple Regressionsmodelle berechnet.

Ergebnisse

LiVs schätzen ihre *individuelle Gesundheitskompetenz* am höchsten in der Dimension *Umgang mit Gesundheitsinformationen* (M=3.21, SD=.51) und am niedrigsten in der *Selbstregulation* (M=2.45, SD=.56). Die latente Profilanalyse bestätigt die vier beruflichen *Selbstregulationstypen* (35.4% Gesundheitstyp, 22.9% Schontyp, 19.2% Typ A, 22.6% Typ B). Im Hinblick auf die dritte Fragestellung konnte gezeigt werden, dass die Gesundheitskompetenzdimensionen *Selbstregulation* und *Selbstkontrolle* hauptsächlich mit der *beruflichen Selbstregulation* zusammenhängen (*Selbstreguluation*: n^2=0.27; *Selbstkontrolle*: n^2=0.15). Unabhängig voneinander hängt der *Gesundheitstyp* (β=.40), *Schontyp* (β=.16) und *Typ A* (β=.12) sowie die Gesundheitskompetenzdimensionen *Selbstregulation (*β=0.39), *Selbstkontrolle (*β=0.15) und *Verantwortungsübernahme* (β=0.21) mit dem Wohlbefinden zusammen. Werden beide Ressourcen zusammen betrachtet, erklären die Dimensionen der *individuellen Gesundheitskompetenz* (*Selbstregulation*: β=0.38; *Selbstkontrolle*: β=0.10; *Verantwortungsübernahme*: β=0.21) mehr Varianz im Wohlbefinden als die *Selbstregulationstypen (Gesundheitstyp*: β=.14).

Diskussion

Aus den Ergebnissen können praktische Implikationen im Hinblick auf die Gesundheitsförderung von angehenden Lehrkräften abgeleitet werden. Erstens könnte eine Lehrkräftebildung, die Aspekte der Selbstregulationsfähigkeit einbezieht, nicht nur das Wohlbefinden von angehenden Lehrkräften verbessern, sondern auch ihre *individuelle Gesundheitskompetenz*. Zweitens könnte eine Förderung der *beruflichen Selbstregulation* durch Lernangebote auch mit der Stärkung der *individuellen Gesundheitskompetenz* von angehenden Lehrkräften verbunden sein. Gleichzeitig wird der Forderung des *Nationalen Aktionsplans Gesundheitskompetenz* nachgegangen, die (Stärkung der) *individuellen Gesundheitskompetenz* im Kontext der Lehrkräftebildung zu verankern (Schaeffer et al., 2018). Im Hinblick auf Forschungsimplikationen sollten zukünftige Studien näher auf die kausalen Zusammenhänge eingehen sowie Kontextfaktoren und/oder Kontrollvariablen miteinbeziehen um ein unfassenderes Bild von möglichen Mediatoren oder Moderatoren des subjektiven Wohlbefindens abbilden zu können.

Zufrieden und hoch effektiv? Eine Meta-Analyse zum Zusammenhang zwischen Berufszufriedenheit von Lehrkräften und ihrer Unterrichtsqualität, Schüleroutcomes, Fehlzeiten und Berufswechselintention

Gyde Wartenberg[1], Karen Aldrup[1], Simon Grund[2], Uta Klusmann[1]
[1]IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany; [2]Universität Hamburg

Berufszufriedenheit stellt neben Burnout einen zentralen Indikator für das berufliche Wohlbefinden dar, das bei Lehrkräften eine bedeutsame Rolle spielt (Jennings & Greenberg, 2009; Wartenberg et al., submitted). Zum einen können Lehrkräfte nicht so leicht in einen anderen Berufskontext wechseln, wenn sie weniger zufrieden sind. Zum anderen tragen Lehrkräfte eine große Verantwortung für die kognitiv-motivationale Entwicklung ihrer Schülerinnen und Schüler, wofür insbesondere die Lehrer-

Schüler Interaktion eine relevante Rolle spielt (Hamre & Pianta, 2001; Tymms et al., 2018). Die Berufszufriedenheit stellt ein zentrales Merkmal dar, das mit der beruflichen Performanz assoziiert ist (Spector, 2022; Weiss, 2002; Wright et al., 2007). Ziel dieser Meta-Analyse ist es daher, die Befunde zum Zusammenhang zwischen der Berufszufriedenheit von Lehrkräften und ihrer beruflichen Performanz im Hinblick auf ihre Unterrichtsqualität, die motivational-kognitive Entwicklung ihrer Schülerinnen und Schüler, Berufswechselintentionen und Fehlzeiten zusammenzufassen.

Hinsichtlich der Konzeptualisierung werden zwei Perspektiven unterschieden, die Berufszufriedenheit zum einen als kognitive Bewertung der Arbeit und zum anderen als affektive Reaktion auf die Arbeit definieren (Aziri, 2011). Entsprechend werden ein kognitionsgesteuerter und ein affektgesteuerter Pfad zur Erklärung des aus Berufszufriedenheit resultierenden Verhaltens postuliert. Aus Sicht der kognitiven Perspektive (Weiss, 2002) sollte eine positive Bewertung der eigenen Arbeit zu aufrechterhaltendem Verhalten, wie beispielsweise der Investition individueller Ressourcen, erhöhtem Engagement, und im Kontext der Lehrkräfte beispielsweise eine höhere Investition von Zeit und Anstrengung in die Unterrichtsvorbereitung und Interaktion mit den Schülerinnen und Schülern. Ergänzend dazu kann die Berufszufriedenheit auch über das affektive Erleben das individuelle Verhalten beeinflussen. Aus der affektiven Perspektive geht eine positive Bewertung der Arbeit mit einem positiven Affekt einher, wodurch das Gedanken-Aktionspotenzial erweitert wird (Fredrickson, 2001; Frenzel et al., 2009; Lazarus & Folkman, 1984; Scherer & Moors, 2019).

Eine umfassende Literatursuche in PsychINFO, Web of Science und Open Grey ergab insgesamt 856 Treffer, von denen 91 Arbeiten in die Analysen eingeschlossen wurde. Zur Beantwortung der Forschungsfrage wurden univariate Meta-Analysen mit Random-Effects Modellen mit dem R-Paket metafor durchgeführt (Field, 2005; Viechtbauer, 2010). Um zu untersuchen, inwieweit die Heterogenität in den Effektstärken auf Studien- und Stichprobenmerkmale zurückgeführt werden kann, wurden Meta-Regressionsmodelle geschätzt.

Die Ergebnisse deuten darauf hin, dass Lehrkräften, die mit ihrem Beruf zufrieden sind, weniger über einen Berufswechsel nachdenken (r = −.44, CI_r = [–0.50, –0.38]) und seltener fehlen (r = −.17, CI_r = [–0.24, –0.10]). Außerdem zeigten sich kleine bis mittlere positive Zusammenhänge zwischen der Berufszufriedenheit von Lehrkräften und ihrer emotionalen Unterstützung (r = .12, CI_r = [0.05, 0.18]), ihrer Klassenführung (r = .18, CI_r = [0.10, 0.26]) und ihrer Instruktionsqualität (r = .23, CI_r = [0.19, 0.27]). Ein gleiches Bild ergab sich für den Zusammenhang zwischen der Berufszufriedenheit von Lehrkräften und der Motivation (r = .29, CI_r = [0.15, 0.42]) sowie Leistung (r = .22, CI_r = [0.05, 0.39]) der Schülerinnen und Schüler. Meta-Regressionen verdeutlichten, dass der negative Zusammenhang zwischen der Berufszufriedenheit von Lehrkräften, ihrer Unterrichtsqualität und der Motivation der Lernenden zwar deutlicher für den Selbstbericht der Lehrkraft war, aber auch für den Fremdbericht bestehen blieb.

Auf Grundlage der Forschungssynthese lassen sich die folgenden Schlussfolgerungen für Forschung und Praxis ableiten: Erstens untersuchten wenige Studien den Zusammenhang zwischen Berufszufriedenheit, emotionaler Unterstützung, Schülerleistungen und -motivation. Hier ist weitere Forschung erforderlich, um bisherige Ergebnisse zu prüfen. Zweitens verdeutlichen die Ergebnisse

der Meta-Regressionsanalyse, dass zukünftige Studien häufiger objektivere Maße implementieren sollten, um eine Überschätzung des Zusammenhangs durch common method bias zu vermeiden. Dass sich auch in Studien, die Outcomes über den Fremdbericht erfassten, signifikante Zusammenhänge zeigten, bestärkt allerdings die Relevanz von Berufszufriedenheit im Lehrkräfteberuf. Auch wenn sich die Kausalität anhand unserer Analysen nicht feststellen lässt, unterstreichen die Ergebnisse den Bedarf an wirksamen Maßnahmen, die die Berufszufriedenheit von Lehrkräften fördern und dadurch ihre berufliche Performanz stärken.

Symposien

F-10: Einfluss von Lehrkraftvariablen auf Schüler:innen in heterogenen Lerngruppen bzw. in inklusiven Klassen

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S05 R03 H20

Einfluss von Lehrkraftvariablen auf Schüler:innen in heterogenen Lerngruppen bzw. in inklusiven Klassen

Chair(s): **Susanne Schnepel** (Universität Zürich, Schweiz)
*Diskutant*in(nen):* **Jasmin Decristan** (Universität Wuppertal)

Die bestmögliche Förderung im inklusiven Unterricht zielt nicht nur darauf, dass alle Lernenden ihren individuellen Voraussetzungen entsprechenden Lernzuwachs erreichen, sondern auch darauf, dass sich alle Lernenden wohl und sozial integriert fühlen, unabhängig von ihrer Leistung oder eines Förderbedarfs. Mehrere Studien zeigen, dass Schüler:innen mit Förderbedarf (z.B. im Lernen oder Verhalten) ein besonderes Risiko haben, sozial abgelehnt zu werden (z.B. Huber & Wilbert, 2012; Wentzel et al., 2021). Diese Ablehnung wird mitunter auch von den Lehrkräften und ihren Erwartungen (z.B. Wang et al., 2018), dem Feedbackverhalten (z.B. Hendrickx et al., 2017) oder der Klassenführungskompetenz (z.B. Garrote et al., 2020) beeinflusst.

Häufig wird auch davon ausgegangen, dass für gelingenden inklusiven Unterricht eine positive Einstellung zur Inklusion (Hellmich & Görel, 2014) und die Zusammenarbeit der Lehrkräfte (z.B. Lütje-Klose & Urban, 2014) eine zentrale Bedingung sind. Allerdings gibt es erst wenig Erkenntnisse zur Wirkung dieser Faktoren (z.B. Avramidis et al., 2019; Sharma et al., 2019).

Was genau eine bestmögliche Förderung in heterogenen Lerngruppen bzw. in inklusiven Klassen ausmacht, konnte empirisch noch nicht ausreichend belegt werden. Das hängt auch damit zusammen, dass die Erforschung dieses Unterrichts äußerst komplex ist, da viele verschiedene Variablen seitens der Lernenden (z.B. Leistung, Verhalten, soziale Akzeptanz), der Lehrkräfte (Einstellung, Erwartungen, Kompetenzen) sowie Einflussfaktoren des Settings (z.B. Anwesenheit einer zweiten Lehrkraft, Kooperation) zusammenspielen. Kaum berücksichtigt wird in den Studien, dass im inklusiven Unterricht oft zwei Lehrkräfte – die Regellehrkraft (RLK) und die sonderpädagogische Lehrkraft (SLK) mit unterschiedlichem Hintergrund, unterschiedlichem professionellem Wissen und unterschiedlichen Einstellungen – anwesend sind, die beide auf das Unterrichtsgeschehen Einfluss nehmen. Jones und Brownell (2014) bezeichnen diese Situation als «nested instruction». Hier ist von Interesse, ob sich zwischen diesen Gruppen von Lehrkräften Unterschiede zeigen und ob deren Wirkmechanismen auf die Schüler:innen unterschiedlich sind.

Das Symposium befasst sich mit der Frage, welchen Einfluss verschiedene Lehrkraftvariablen auf die Leistungsentwicklung, die soziale Integration und das schulische Wohlbefinden der Kinder haben. Um mehr über die Wirkmechanismen im Rahmen der «nested instruction» herauszufinden, wird in Studien, an denen sowohl RLK als auch SLK teilgenommen haben, zwischen RLK- und SLK-Variablen differenziert.

In den ersten beiden Beiträgen wird der Einfluss von Lehrkraftvariablen auf die soziale Akzeptanz unter Berücksichtigung von besonderem Förderbedarf (niedrige Schulleistungen in Beitrag 1, Verhaltensprobleme in Beitrag 2) untersucht.

Der dritte und vierte Beitrag stellen sich der Herausforderung, die sogenannte «nested instruction», den Einfluss von gemeinsam unterrichtenden Lehrkräften, auf die Schüler:innen (Outcome Variable Leistung und Wahrnehmung der unterstützenden Lernbeziehung) zu untersuchen.

Beiträge des Symposiums

Verhaltensprobleme und soziale Integration – welchen Einfluss haben die Classroom-Management-Kompetenzen von Lehrkräften im inklusiven Unterricht?

Christian Huber[1], Markus Spilles[1], Philipp Nicolay[1], Thomas Hennemann[2], Michael Grosche[1], Johannes König[2]
[1]Universität Wuppertal, [2]Universität zu Köln

Die soziale Integration von Schulkindern in der Schule ist ein psychologisches Grundbedürfnis (Ryan & Deci, 2001) und gilt als Grundvoraussetzung für schulisches Lernen und eine gute emotional-soziale Entwicklung. Soziale Ausgrenzung geht hingegen mit einer Vielzahl schulisch relevanter Risikofaktoren wie Aggressivität, Angststörungen, depressive Symptome und schwachen Schulleistungen oder Schulabsentismus einher (Newcomb, Bukowski & Pattee, 1993; Wentzel, Jablansky & Scalise, 2021; Flook et al., 2005). Je nach Studie sind Kinder mit Verhaltensproblemen bis zu dreimal häufiger von sozialer Ausgrenzung im inklusiven Unterricht betroffen als Ihre Klassenkameraden ohne besonderen Förderbedarf (Bless, 2007; Huber, Nicolay & Schulze, 2021).

Bislang befasst sich nur ein geringer Anteil der Studien mit der Frage, wie man soziale Integrationsprozesse in der Schule evidenzbasiert beeinflussen kann. Diskutiert werden im Wesentlichen vier Ansätze (Huber, 2019), von denen wir uns hier auf zwei beziehen. Das social skills deficit model (Asher, Renshaw & Hymel, 1982) geht davon aus, dass soziale Integration durch eine Förderung der sozialen Kompetenzen der ausgegrenzten Kinder verbessert werden kann. Die social referencing theory (Feinman, 1992) geht davon aus, dass öffentliches Lehrkraftfeedback die soziale Integration beeinflusst. In der vorliegenden Studie prüfen wir, ob beide Ansätze mit Unterschieden im Classroom-Management der Lehrkräfte zusammenhängen.

Fragestellung

Die eingereichte Studie verbindet beide o.g. Ansätze in einem Wirkmodell und geht davon aus, dass Classroom-Management einerseits das Sozialverhalten von Schulkindern mit Verhaltensproblemen günstig beeinflusst und andererseits die Häufigkeit von reaktivem negativem Lehrkraftfeedback reduziert. Es wird vorhergesagt, dass der Zusammenhang zwischen sozialer Integration und Verhaltensproblemen im Unterricht einerseits insbesondere in Schulkassen hoch ist, in denen die Lehrkräfte ein schwaches Classroom-Management (CRM) zeigen. Andererseits müsste sich dieser Zusammenhang reduzieren, wenn Lehrkräfte über gute Classroom-Management-Kompetenzen verfügen.

Methode

Die Hypothese wurde in einem Teilprojekt einer Querschnittsstudie mit 1112 Schulkinder (Klasse 3-4) aus 61 Klassen untersucht. Als abhängige Variable (AV) wurde die eigene soziale Integration aus Sicht der Kinder durch die soziometrische Methode (Moreno, 1974) erhoben. Berechnet wurden Wahlstatus (AV1: WST) und Ablehnungsstatus (AV2: AST). Als unabhängige Variablen wurden die Verhaltensprobleme pro Schüler*in durch eine fünfstufige Likertskala (UV1) und die selbst

eingeschätzen Classroom-Management-Kompetenzen durch drei in Vorstudien überprüfte CRM-Skalen (UV2: Routinen; UV3: Regeln; UV4: Störungsausmaß) jeweils aus Lehrkraftsicht erhoben.

Ergebnisse

Die Datenanalyse erfolgt durch zwei Mehrebenenregressionsanalysen. Für die Datenanalyse wurde für jede AV ein Random-Intercept-Modell berechnet, in dem die individuellen Schülermerkmale auf Level 1 (AV1-2, UV1) und die drei lehrkraftbezogenen Classroom-Management-Kompetenzen (UV2-4) auf Level 2 angesiedelt waren. Erwartet wurde jeweils eine Cross-Level-Interaktion der Classroom-Management-Kompetenzen auf den Zusammenhang zwischen Verhaltensproblemen und sozialer Integration.

Die Ergebnisse zeigten den bereits aus der Literatur bekannten negativen signifikanten Effekt (b = -16.58) der Verhaltensprobleme auf den WST sowie einen positiven Effekt (β = 13.68) der Verhaltensprobleme auf den AST. Für den WST zeigte sich eine signifikante Cross-Level-Interaktion lediglich für die CRM-Skala Routinen (b = 6.15), jedoch entgegengesetzt zur vorhergesagten Richtung. Für den AST waren die Cross-Level-Interaktionen für alle drei CRM-Skalen signifikant, jedoch waren auch hier die Effekte für zwei Skalen (Routinen, Störungen) entgegengesetzt der vorhergesagten Richtung. Lediglich die Cross-Level-Interaktion zwischen der CRM-Skala Regeln und dem AST zeigte sich in der vorhergesagten Richtung (b = -2.33).

Diskussion

Die Ergebnisse werden so interpretiert, dass unterschiedliche Classroom-Management-Dimensionen den Zusammenhang zwischen Verhaltensproblemen unterschiedlich beeinflussen und somit im Unterricht auch nicht in die gleiche Richtung wirken könnten. Classroom-Management mit zahlreichen Routinen und wenigen Unterrichtsstörungen stärkt den Zusammenhang zwischen Verhaltensproblemen und sozialer Ausgrenzung – möglicherweise weil Verhaltensprobleme bei effektiven Unterrichtsroutinen häufiger auffallen und Lehrkräfte die betroffenen Kinder verbal regulieren müssen. Andersherum zeigte sich für stark regelbasiertes und eher präventiv ausgerichtetes CRM ein abschwächender Effekt auf den Zusammenhang zwischen Verhaltensprobleme und sozialer Ausgrenzung – möglicherweise weil sich hierdurch Unterrichtsstörungen reduzieren und Lehrkräfte somit weniger verbal regulieren müssen.

Der Einfluss von Lehrkraftvariablen auf das Leseverständnis von Grundschüler:innen inklusiven Klassen

Susanne Schnepel, Maria Wehren Müller, Meret Stöckli, Simon Luger, Elisabeth Moser Opitz
Universität Zürich

Theoretischer Hintergrund

Studien zeigen, dass die Kompetenzen der Lehrkräfte für die Entwicklung der Lernenden bedeutsam sind (Baumert et al., 2010; Hill et al., 2008; Blömeke et al., 2022). Jedoch ist kaum erforscht, ob und wie das professionelle Wissen der Lehrkräfte zur Leseförderung den Leistungszuwachs der Schüler:innen im inklusiven Unterricht beeinflusst.

Häufig wird angenommen, dass die Einstellung zur Inklusion wichtig ist für gelingenden inklusiven Unterricht (de Boer et al., 2011; Avramidis & Norwich, 2002) und es konnte ein Zusammenhang der

Einstellung zur Inklusion mit der wahrgenommenen Selbstwirksamkeit nachgewiesen werden (Hellmich & Görel, 2014). Ausgehend davon wird angenommen, dass ein hohes professionelles Wissen die Selbstwirksamkeit erhöht und somit indirekt zu einer positiveren Einstellung zur Inklusion führt. Ob sich dies auf die Schüler:innen auswirkt, wurde noch nicht untersucht.

Weiter gilt eine gute Zusammenarbeit von Regellehrkraft (RLK) und sonderpädagogischer Lehrkraft (SLK) als Bedingung für gelingenden inklusiven Unterricht (Grosche et al., 2020; Lütje-Klose & Urban, 2014). Bislang wurde die Wirkung der Kooperation allerdings kaum untersucht, insbesondere auch hinsichtlich der Auswirkung auf die Schüler:innen.

Schließlich muss berücksichtigt werden, dass im inklusiven Unterricht die sogenannte „nested instruction“ (Jones & Brownell, 2014), die gleichzeitige Einflussnahme von RLK und SLK auf die Schüler:innen, eine zentrale Rolle spielt und deshalb beide Lehrkräfte berücksichtigt werden müssen.

Fragestellung

• Welchen Einfluss haben das professionelle Wissen zur Leseförderung, die Einstellung zur Inklusion und die Kooperationszufriedenheit von der RLK und SLK auf die Leseleistung von Grundschüler:innen?

• Wie hängen die Einstellung zur Inklusion und Kooperationszufriedenheit von RLK und SLK zusammen?

Methode

Stichprobe: 70 inklusive Klassen (2.-4. Klasse) mit Teams aus RLK und SLK und 1407 Kinder.

Erhebungen Kinder:

Leseverständnis ELFE-Test (Lenhard et al., 2017) zu Beginn (t1) und am Ende des Schuljahres (t2).

Allgemeine Denkfähigkeit: CFT 1-R (Weiß & Osterland, 2013) bzw. CFT 20-R (Weiß, 2006).

Lehrkraftvariablen:

Professionelles Wissen Leseförderung: selbstentwickeltes Instrument zum Rating von Interviews zum Thema Leseförderung (vierstufiges Item, Übereinstimmung Rater:innen g = .87)

Einstellung zur Inklusion: Fragebogen Attitudes towards the inclusion of students with SEN (AIS) (Sharma & Jacobs, 2016, 8 Items, αRLK = .83, αSLK = .63)

Kooperationszufriedenheit: Fragebogen zur Arbeit im Team (FAT) (Gebhard et al., 2014, 10 Items, αRLK = .84, αSLK = .83)

Zur Analyse der Zusammenhänge und der Effekte auf die Leseleistung der Kinder zu t2 wurden Mehrebenen-Pfadmodelle gerechnet.

Ebene 1: Kontrollvariablen allgemeine Denkfähigkeit, Geschlecht und Leseleistung zu t1.

Auf Ebene 2: professionelles Wissen Leseförderung, Einstellung zur Inklusion und Kooperationszufriedenheit jeweils von der RLK und SLK einzeln als Prädiktoren. Es wurde davon ausgegangen wurde, dass das Professionswissen direkt und mediiert über die Einstellung auf die Leseleistung wirkt. Zudem wurde angenommen, dass es inter- und intraindividuelle Zusammenhänge zwischen Einstellung und Kooperationszufriedenheit gibt.

Ergebnisse und ihre Bedeutung

Das Professionswissen der SLK hat einen direkten Einfluss auf die Leseleistung der Kinder (β=.33, p=.018) und auf die Einstellung (β=.33, p=.003), das Professionswissen der RLK hat keinen signifikanten Effekt hat. Auch die Einstellung und die Kooperationszufriedenheit von RLK und SLK haben keinen Effekt auf die Leseleistung. Bei den RLK korrelieren Einstellung und Kooperationszufriedenheit (r=.30, p=.006) und die Kooperationszufriedenheit von RLK und SLK korrelieren (r=.49, p<.001).

Die Ergebnisse zeigen, dass die Lehrkraftvariablen von RLK und SLK differenziert betrachtet werden müssen, da sie das Lernen der Kinder unterschiedlich beeinflussen. So ist das Professionswissen der SLK für die Leseentwicklung bedeutsam und steht in einem Zusammenhang zur Einstellung zur Inklusion. Allerdings ist die Richtung dieses Zusammenhangs nicht klar und hier müssten weitere Untersuchungen folgen. Hinsichtlich des Einflusses von Einstellung zur Inklusion und der Lehrkräftekooperation zeigt sich, dass sich dieser auf der Ebene der Lehrkräfte zu zeigen scheint. Es müsste somit in Zukunft genauer spezifiziert werden, für welche Aspekte gelingenden inklusiven Unterrichts die beiden Variablen bedeutsam sind.

Der Einfluss von Lehrkraftvariablen auf das Wohlbefinden von Grundschülerinnen und -schülern in inklusiven Klassen

Simon Luger, Susanne Schnepel, Maria Wehren Müller, Meret Stöckli, Elisabeth Moser Opitz
Universität Zürich

Theoretischer Hintergrund

Für die erfolgreiche Umsetzung inklusiven Unterrichts werden verschiedene Lehrkraftvariablen als wichtig erachtet. Dazu gehört u. a. die Einstellung zur Inklusion von Schüler:innen mit besonderem Bildungsbedarf (bB). Positiv eingestellte Lehrkräfte haben ein höheres Maß an Selbstwirksamkeit für inklusive Praktiken und setzen diese daher häufiger um (Avramidis et al., 2019). Auch wird angenommen, dass gemeinsames Lernen (Unterrichtung aller Schüler:innen im gleichen Raum) besser gelingt, wenn die Regellehrkräfte (RLK) und Sonderschullehrkräfte (SLK) zufrieden sind mit der Kooperation im Team (Gebhard et al., 2014). Zudem wird vermutet, dass die Intensität der Kooperation ausserhalb der Unterrichtszeit (Austausch, Arbeitsteilung, Kokonstruktion) wichtig ist (Grosche, 2020). Studien zur Wirksamkeit entsprechender Lehrkraftvariablen untersuchen in der Regel Zusammenhänge mit der Leistung von Schüler:innen (z. B. Dubberke et al., 2008), kaum jedoch mit sozio-emotionalen Aspekten, wie etwa der Beziehung zur Lehrkraft oder dem unterstützenden Klima. Untersuchungen zeigen, dass insbesondere Schüler:innen mit bB in inklusiven Klassen sozio-emotionalen Risiken ausgesetzt sind (Kullmann, 2015). Es ist deshalb wichtig, dass die Schüler:innen in inklusiven Klassen von einem protektiven Klassenklima und einer von Wertschätzung und Unterstützung geprägten Beziehung zur Lehrkraft getragen werden (Kröske, 2020). Diese Studie untersucht im Kontext inklusiven Unterrichts die Wirkung verschiedener Lehrkraftvariablen auf die Veränderung der Wahrnehmung der unterstützenden Beziehung der Schüler:innen zur Lehrkraft. Bei den Lehrkraftvariablen stehen die Einstellung zur Inklusion, die Kooperationszufriedenheit der RLK und SLK, die Umsetzung gemeinsamer Lernsituationen (UgL) und die Kooperationsintensität im Fokus.

Fragestellung

Inwiefern mediieren die Umsetzung gemeinsamer Lernsituationen und die Kooperationintensität den Zusammenhang zwischen der Einstellung der Lehrkräfte (RLK, SLK) zur Inklusion bzw. der Kooperationszufriedenheit und der Veränderung der von den Schüler:innen wahrgenommenen unterstützenden Beziehung zu den Lehrkräften?

Methode

Die Stichprobe umfasst 76 inklusive Klassen (2.-4. Jahrgang) mit Teams aus RLK und SLK und 992 Schüler:innen. Die unterstützende Beziehung wurde lehrkraftspezifisch (RLK und SLK) zu zwei Messzeitpunkten (t1 und t2) mit einer Skala in Anlehnung an Fauth et al., 2014 erhoben (4 Items, $\alpha_{RLK} = .69$, $\alpha_{SLK} = .80$).

Zur Erfassung der Lehrkraftvariablen wurden folgende Instrumente eingesetzt:

Einstellung zur Inklusion: Fragebogen Attitudes towards the inclusion of students with SEN (AIS) (Sharma & Jacobs, 2016, 8 Items, $\alpha_{RLK} = .83$, $\alpha_{SLK} = .63$)

Kooperationszufriedenheit: Fragebogen zur Arbeit im Team (FAT) (Gebhard et al., 2014, 10 Items, $\alpha_{RLK} = .84$, $\alpha_{SLK} = .83$)

Umsetzung gemeinsamer Lernsituationen und Kooperationsintensität: Rating von strukturierten Einzelinterviews mit RLK und SLK. Beim Rating der UgL wurde eingeschätzt, wie häufig Schüler:innen mit bB im selben Raum mit allen anderen unterrichtet werden (1 = sehr seperativ bis 4 = sehr inklusiv). Beim Rating der Kooperationsintensität wurde eingeschätzt, wie intensiv die gemeinsame Unterrichtsplanung und -reflexion erfolgt (1 = nur Informationsaustauch, 2 = arbeitsteilige Kooperation, 3 = kokonstruktive Kooperation). Die Rater:innenübereinstimmung lag zwischen $\kappa = .42$ und .77.

Analyse: Je ein Mehrebenen-Pfadmodelle für die RLK und die SLK (Ebene 1: Kontrollvariablen Geschlecht, Alter, IQ und Wahrnehmung unterstützende Lernbeziehung t1. Ebene 2: Einstellung zur Inklusion, Kooperationszufriedenheit, UgL, Kooperationsintensität.

Annahme: Die Wirkung der Einstellung und Kooperationszufriedenheit auf die unterstützende Beziehung wird durch die UgL und die Kooperationsintensität mediiert.

Ergebnisse und Bedeutung

Bei den SLK haben die Einstellung und die Kooperationszufriedenheit einen positiven Einfluss auf die UgL und die Kooperationsintensität, was sich auch auf die Veränderung der von den Schüler:innen wahrgenommenen unterstützenden Beziehung zur Lehrkraft auswirkt ($\beta = .19$, $p = .041$). Bei den RLK hat die Einstellung keinen Einfluss auf die UgL und die Kooperationsintensität und es zeigen sich auch keine Auswirkungen auf die Schüler:innen. Die Ergebnisse weisen darauf hin, dass sowohl die Einstellung als auch die Zusammenarbeitszufriedenheit der SLK bedeutsam sind für die Umsetzung von gemeinsamen Lernsituationen und die Zusammenarbeitsintensität, mit einer positiven Auswirkung auf die Lernenden. Im Vortrag werden mögliche (methodische und Inhaltliche) Erklärungen dieses Ergebnisses diskutiert.

F-11: Adaptive Lernunterstützung im mathematischen, sprachlichen und kognitiven Förderbereich in der Kita

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* T03 R02 D26

Adaptive Lernunterstützung im mathematischen, sprachlichen und kognitiven Förderbereich in der Kita

Chair(s): **Manfred Holodynski** (Universität Münster, Germany)
*Diskutant*in(nen):* **Kristin Krajewski** (Pädagogische Hochschule Ludwigsburg)

Der gemeinsame Gegenstand aller Symposiumsbeiträge ist die adaptive Lernunterstützung von Kindern zwischen drei und sechs Jahren durch einen Erwachsenen während gemeinsamer spielerischer und/oder erzählerischer Interaktionen, wie sie sich auf die Fähigkeitsentwicklung der Kinder auswirken und durch Fortbildungsmaßnahmen gefördert werden können. Zwei Beiträge fokussieren auf den mathematischen Bildungsbereich, einer auf den sprachlichen und der vierte auf den kognitiven Bereich.

Der Beitrag „Adaptive Gestaltung lernunterstützender, kognitiv aktivierender Fachkraft-Kind-Interaktionen in der Kita“ umfasst die Analyse von Fachkraft-Kind-Interaktionen in Kitas unter der Perspektive, wie kognitiv aktivierend Fachkräfte diese Interaktionen gestalten und inwiefern dies davon abhängt, ob eine Fachkraft mit einem Kind interagiert, das aus ihrer Sicht sozial benachteiligt ist oder nicht. Die videobasierten Beobachtungen ergaben, dass die Interaktionen nur wenig kognitiv aktivierend waren, wobei diese gegenüber vermeintlich sozial benachteiligten Kindern anregender waren als gegenüber nicht sozial benachteiligten.

Der Beitrag „Die Wirksamkeit adaptiver Sprachförderung nach den Prinzipien des Dialogischen Lesens“ adressiert die adaptive Sprachförderung mit Hilfe des Dialogischen Lesens in der Kita. Vier Evaluationsstudien werden mit folgenden Ergebnissen referiert: (1) Mit Hilfe des Dialogischen Lesens konnten auch Kinder im Zweitspracherwerb kurzfristig gefördert werden. (2) Auch gegenüber zwei alternativen Sprachförderprogramm zeigten sich positive Effekte, wobei diese von der Adaptivität des Dialogischen Lesens abhingen. (3) Eine optimierte Fortbildung von Fachkräften im Einsatz Dialogischen Lesens erbrachte auch langfristig kindliche Fördererfolge. (4) Das Dialogische Lesen war auch unter alltagspraktischen Kita-Bedingungen erfolgreich.

Der Beitrag „Mathematischer Sprachinput in Krippen“ behandelt eine spezifische Fördermethode für mathematische Basisfähigkeiten, den Math Talk, bei dem sich Kita-Fachkräfte mit ihren Kindern über mathematische Phänomene und Themen in adaptiver, kindgerechter Weise unterhalten. Berichtet wird die Evaluation einer Fortbildung mit Kita-Fachkräften, inwiefern sich deren Math Talk fördern lässt. Die Effekte dieser Fortbildung werden mit einer Fortbildung zu globalen Interaktionen und mit einer unbehandelten Kontrollgruppe verglichen. Dabei zeigte sich ein stärkerer Zuwachs des Math Talk in der Fortbildungs- gegenüber der Kontrollgruppe.

Im Beitrag „Adaptive Lernunterstützung in einer regelspielbasierten Förderung mathematischer Basisfähigkeiten Vier- bis Sechsjähriger“ wird die Evaluation einer Förderung mathematischer Basisfähigkeiten von Vorschulkinder berichtet, bei der eine Gruppe an Spielleitungen spielbegleitende videobasierte Coachings zur Optimierung ihrer adaptiven Lernunterstützung erhielten, während die andere Gruppe keine Coachings bekam. Beide Fördergruppen konnten die mathematischen Basisfähigkeiten ihrer Kinder signifikant verbessern im Vergleich zu einer unbehandelten

Kontrollgruppe. Allerdings waren die gecoachten Spielleitungen in der Förderung des kindlichen mathematikbezogenen Regelspielverständnisses erfolgreicher als die nicht gecoachten Spielleitungen.

Die Beiträge werden von der eingeladenen Diskutantin auf ihren Beitrag zur adaptiven Lernunterstützung in der Kita diskutiert.

Beiträge des Symposiums

Adaptive Gestaltung lernunterstützender, kognitiv aktivierender Fachkraft-Kind-Interaktionen in der Kita

Heike Wadepohl, Lisa Keller, Theresa Johannsen, Katja Mackowiak
Universität Hannover

Eine Vielzahl (inter-)nationaler Studien verweist auf die Bedeutung der Qualität frühkindlicher Institutionen im Hinblick auf kindliche Entwicklungsoutcomes sowie eine Verringerung von Bildungsdisparitäten. Die Gestaltung von Fachkraft-Kind-Interaktionen im Kita-Alltag gilt dabei als ein zentraler Indikator pädagogischer Qualität (Perlman et al., 2016; Ulferts et al., 2019).

Während die Fachkräfte ihre Interaktionen mit den Kindern bereits zur Weiterentwicklung wichtiger kindlicher Kompetenzen nutzen (z.B. sozial-emotionale, sprachliche; vgl. Betz et al., 2016; Wadepohl, 2016), scheint insbesondere die Gestaltung lernunterstützender Interaktionen, die das Denken und Problemösen von Kindern anregen, eine Herausforderung darzustellen. Insbesondere kognitiv aktivierende Impulse sind dabei im Kita-Alltag nur selten bzw. in geringer Qualität zu beobachten (vgl. zusammenfassend Mackowiak et al., 2022).

Als wesentliche Gelingensbedingung einer alltagsintegrierten (kognitiven) Förderung wird das Maß an Adaptivität, also eine möglichst optimale Passung zwischen den Anregungen der Fachkraft, den individuellen Voraussetzungen und Interessen der Kinder und dem Lernangebot, herausgestellt (Baltruschat, 2018; Lachner et al., 2018; Neuß, 2014). Allerdings ist das Konzept der Adaptivität insbesondere für den vorschulischen Bereich bisher kaum ausdifferenziert sowie - insbesondere durch die in der quantitativen Forschung gängige gruppenbezogene Erhebungsperspektive - empirisch nur in Ansätzen erfassbar (Hardy et al., 2019).

Im Rahmen dieses Beitrags sollen daher anhand von Daten aus den vom Niedersächsischen Ministerium für Wissenschaft und Kultur geförderten KoAkiK-Projekten zur „Kognitiven Aktivierung in (inklusiven) Kitas“ erste Ideen zur adaptiven Gestaltung lernunterstützender, kognitiv aktivierender Fachkraft-Kind-Interaktionen vorgestellt und diskutiert werden. Es stehen folgende Fragen im Zentrum des Vortrags:

(1) Wie (gut) gestalten frühpädagogische Fachkräfte lernunterstützende, kognitiv aktivierende Interaktionen mit den Kindern?

(2) Welche Unterschiede in der Lernunterstützung lassen sich zwischen den beobachteten Interaktionen feststellen und inwiefern können daraus erste gruppenbezogene Kriterien für Adaptivität abgeleitet werden?

(3) Zeigen sich Unterschiede in der Lernunterstützung der Fachkräfte unter Berücksichtigung des kognitiven Entwicklungsstandes des mit der Fachkraft interagierenden Kindes? Welche Implikationen hat diese individualisierte Perspektive auf die Erfassung von Adaptivität?

Als Datengrundlage dienen jeweils ca. 20-minütige Videografien von dyadischen Fachkraft-Kind-Interaktionen, die im Hinblick auf die Qualität und Gestaltung der Lernunterstützung mittels zweier Beobachtungsverfahren ausgewertet wurden: mit dem Classroom Assessment Scoring System (CLASS Pre-K, Pianta et al., 2008) und dem mikroanalytischen Beobachtungsinstrument zur Erfassung der Lernunterstützung im Kita-Alltag (B-LUKA, Johannsen & Keller, 2022). Von insgesamt 85 Fachkräften liegt eine ausgewertete dyadische Situation vor; bei 42 Fachkräften konnten zwei dyadische Situationen mit zwei verschiedenen Kindern ausgewertet werden, wobei ein Kind aus Sicht der Fachkraft aus sozial benachteiligten und eines aus sozial nicht benachteiligten Verhältnissen stammt. Bei 41 Dyaden liegen zudem Testdaten zu den kognitiven Kompetenzen der interagierenden Kinder vor.

Erste Ergebnisse weisen darauf hin, dass es den Fachkräften nur selten bzw. in geringer bis mittlerer Qualität gelingt, (potenziell) kognitiv aktivierende Interaktionen mit den Kindern zu gestalten. Im Dyaden-Vergleich fallen auf deskriptiver Ebene durchgängig leicht höhere Werte in den Dyaden mit dem vermeintlich sozial benachteiligten Kind auf; die Unterschiede sind in der vorliegenden Stichprobe jedoch nicht signifikant. Unter dem Gesichtspunkt der Adaptivität lässt sich dennoch vorsichtig vermuten, dass die Fachkräfte die dyadischen Situationen insbesondere bei den aus ihrer Sicht benachteiligten Kindern als explizite Bildungschance nutzen, während sie bei den vermeintlich nicht benachteiligten Kindern möglicherweise andere Schwerpunkte setzen und das vorhandene (kognitive Bildungs-)Potenzial nicht ausreichend erkennen bzw. ausschöpfen (vgl. auch Wadepohl & Bethke, 2022). Hier sollen differenzierte (Einzelfall)-Analysen unter Einbezug des kognitiven Entwicklungsstandes der mit der Fachkraft interagierenden Kinder weitere Hinweise auf die Adaptivität der Lernunterstützung aus individualisierten Perspektive geben, die auf der GEBF-Tagung vorgestellt und diskutiert werden.

Keywords: Adaptivität, Lernunterstützung, kognitive Aktivierung, Interaktionsqualität, Kindertageseinrichtung

Die Wirksamkeit adaptiver Sprachförderung nach den Prinzipien des Dialogischen Lesens

Marco Ennemoser, Melanie Besca
PH Ludwigsburg

Obwohl es eine Vielzahl an Ansätzen zur Sprachförderung im Kindergarten gibt, existieren nur wenige Evaluationsstudien zur Wirksamkeit der Förderkonzepte, die darüber hinaus zu überwiegend ernüchternden Ergebnissen kommen (z.B. Schneider, 2018; Egert, Galuschka, Groth, Hasselhorn & Sachse, 2020). Internationale Befunde lassen darauf schließen, dass eine Konzentration auf das Wesentliche, nämlich die sprachförderliche Interaktion, zielführender sein könnte als breiter angelegte Programme zur Sprachförderung. Prototypisch für eine rein interaktionsbasierte Förderung ist der Ansatz des Dialogischen Lesens, der die Vorlesesituation in der Kleingruppe als Plattform nutzt, um Interaktionen mit Kindern zu realisieren. Auch wenn das Dialogische Lesen üblicherweise nicht unter diesem Label geführt wird, handelt es sich dabei um einen konsequent adaptiven Förderansatz.

Durch gezielte Fragetechniken werden die Kinder zur Sprachproduktion angeregt, so dass die eingesetzten Techniken zum korrektiven Feedback bzw. zur Modellierung dieser Äußerungen nahezu zwangsläufig auf dem jeweiligen sprachlichen Kompetenzniveau ansetzen.

International belegen zahlreiche Studien die Wirksamkeit der Methode für verschiedene Zielgruppen und in verschiedenen Kontexten (Flack, Zoe M., Field & Horst, 2018; Mol et al., 2009; Swanson et al., 2011). Der vorliegende Beitrag befasst sich mit der Frage, ob der Ansatz des Dialogischen Lesens auch als wirksame Methode der Sprachförderung in den Kindergartenalltag in deutschen Kindergärten eingebettet werden kann. Evidenz hierfür liefert eine Serie von vier Studien. Eine kleinere Studie mit N = 45 Kindern zeigt, dass der Ansatz auch im Rahmen der Sprachförderung bei Kindern im Zweitspracherwerb kurzfristig positive Effekte auf die Sprachkompetenz nach sich ziehen kann (Ennemoser, Kuhl & Pepouna, 2013). Als Kontrollbedingung diente die herkömmliche Sprachförderung im Rahmen von Vorbereitungskursen für den Schuleintritt. In einer zweiten Studie mit Prä-Posttest-Follow-up Design wurde die Wirksamkeit des Dialogischen Lesens mit zwei alternativen Förderprogrammen verglichen (N1 = 59, N2 = 58 , N3 = 58, N4 = 57) (Ennemoser & Hartung, 2017). Auch hier konnten positive Effekte des Dialogischen Lesens bestätigt werden. Allerdings erwiesen sich die Effekte nicht als nachhaltig und als abhängig von der realisierten Durchführungsqualität. In einer größer angelegten dritten Studie mit insgesamt 374 Kindern wurde die Fortbildung für die Förderkräfte optimiert, wodurch die positiven Effekte auch langfristig abgesichert werden konnten (Ennemoser, Lehnigk, Hohmann & Pepouna, 2015). Eine vierte Studie zielte schließlich auf die Frage ab, ob die Maßnahme auch unter alltagspraktischen Bedingungen und bei einer Durchführung durch pädagogische Fachkräfte der Einrichtungen ihre Wirksamkeit entfalten kann (Besca, Reiner & Ennemoser, in Vorb.). Auch hier ergaben sich signifikante Vorteile der mittels Dialogischem Lesen geförderten Kinder (N = 150) im Vergleich zu einer Kontrollgruppe (N = 66). Insgesamt liefern die Studien vergleichsweise konsistente und in mehreren Studien replizierbare Evidenz für die Wirksamkeit des Dialogischen Lesens, auch unter weniger gut kontrollierten Bedingungen des Kindergartenalltags.

Keywords: Sprachförderung, dialogisches Lesen, Kindergarten, Evaluation

Mathematischer Sprachinput in Krippen

Johanna Klemm[1], Simone Lehrl[1], Anja Linberg[2]
[1]Universität Bamberg, [2]Deutsches Jugendinstitut

Viele Studien belegen Zusammenhänge zwischen der Interaktionsqualität, welche Kinder in Kindertageseinrichtungen mit pädagogischen Fachkräften erleben, und ihren Kompetenzen (Mashburn et al., 2008; Vandell et al., 2016). Gleichzeitig weisen Studien darauf hin, dass Fachkraft-Kind-Interaktionen durch Fortbildungen verbessert werden können (Hamre et al, 2012, Brunsek et al., 2020). Dabei kann zwischen globalen und domänenspezifischen Interaktionen unterschieden werden (Hamre & Pianta, 2007). Der vorliegende Beitrag nimmt domänenspezifisch mathematische Interaktionen in den Fokus, denn bereits im Alter von drei Jahren (Clements & Sarama, 2007; Dubowy, Ebert, von Maurice & Weinert, 2008) und vier Jahren (Klibanoff et al. 2006) bestehen signifikante Unterschiede in den mathematischen Kompetenzen von Kindern. Studien zeigen zudem, dass diese frühen mathematischen Kompetenzen wiederum für die späteren mathematischen

Fähigkeiten und Schulerfolge auch prädizieren, wenn die allgemeine Intelligenz kontrolliert wird (Aunola, Leskinen, Lerkkanen & Nurmi, 2004; Dornheim, 2008; Jordan, Kaplan, Locuniak & Ramineni, 2007; Lee & Burkham, 2002). Daher ist die mathematische Anregung im sehr frühen Alter von Interesse. Studien, welche bislang Kinder im Alter ab vier Jahren untersuchten, konnten zeigen (Klibanoff et al., 2006; Boonen, et al., 2011; Ramani, 2015), dass besonders ein mathematisches Sprachangebot, auch als Math Talk bekannt, eng in Verbindung mit mathematischen Kompetenzen steht (Carey, 2009; Krajewski & Schneider, 2009; Negen & Sarnecka, 2012; Spelke, 2017). Math Talk ist eine Form domänenspezifischer Interaktion und beschreibt mathematischen Sprachinput (Boonen et al., 2011, Klibanoff et al., 2006). Zum einen beinhaltet Math Talk numerische Begriffe und Erkundungen des Zahlensystems (z.B. Zahlwörter, Kardinalität). Zum anderen umfasst das Konzept auch räumliches und quantitatives Vokabular (z.B. unter, auf, mehr, wenige, keine) (Ramani, Rowe, Eason, & Leech, 2015). Bislang ist unklar, wie viel und welche Formen von Math Talk pädagogische Fachkräfte in Krippen anwenden und wie diese Form alltagsintegrierter, mathematischer Förderung gestärkt werden kann. Daher geht der vorliegende Beitrag den Fragen nach (a) welche Arten von Math Talk kommen wie oft in Fachkraft-Kind-Interaktionen in der Krippe vor und (b) kann Math Talk durch eine globale oder mathematik-spezifische Fortbildung verbessert werden. Methode

Zur Beantwortung der Fragen werden Daten des EarlyMath-Projektes genutzt. Einem quasi-experimentellen Design folgend nahmen 45 pädagogische Fachkräfte und 210 Kinder im Alter zwischen 24 und 32 Monaten sowie ihre Eltern an der Studie teil. Die Fachkräfte wurden zufällig einer 2,5 tägigen Fortbildung zu globalen Interaktionen, zu mathematik-spezifischen Interkationen oder einer Kontrollgruppe (welche an einem Informationsnachmittag teilnahmen) zugeordnet. Alle Fachkräfte wurden im Prä- und Posttest in Interaktion mit bis zu drei Kindern in vorstrukturierten Spielsituationen (Buch und Spiel vorgegeben) gefilmt. Die Videosequenzen wurden anschließend transkribiert und auf Äußerungsebene mit einem speziell in Anlehnung an die Arbeiten von Klibanoff et al. (2006), Boonen, et al. (2011) und Ramani et al. (2015) für Math Talk entwickelten Kodierschema kodiert. Gruppenunterschiede zwischen den Interventionsgruppen und der Kontrollgruppe wurden in IBM SPSS 25 mit ANOVAs mit Messwiederholungen analysiert. Ergebnisse Fachkräfte sprechen über vielfältige mathematische Themen, wie z.B. Kardinalität, Reihenfolgen oder Formen. Durchschnittlich sprechen Fachkräfte 7 unterschiedliche mathematische Inhaltsbereiche in einer ca. 20-minütigen Interaktionssituation an. Die Häufigkeit variiert stark zwischen unterschiedlichen Fachkräften und auch zwischen den mathematischen Inhaltsbereichen. Des Weiteren nutzen pädagogische Fachkräfte dazu eine große Anzahl und Diversität mathematischer Begriffe. Erste ANOVA-Analysen der längsschnittlichen Daten weisen auf einen stärkeren Zuwachs mathematischer Sprache in der mathematik-spezifischen Interventionsgruppe hin. Pädagogische Fachkräfte sprechen mit Kindern nach der Fortbildung häufiger über mathematische Themen als Fachkräfte der Kontrollgruppe. Die Befunde werden mit Blick auf die Bedeutung und Förderung domänenspezifischer Lernunterstützung in Kindertagesstätten diskutiert.

Adaptive Lernunterstützung in einer regelspielbasierten Förderung mathematischer Basisfähigkeiten Vier- bis Sechsjähriger

Manfred Holodynski, Dorothee Seeger, Elisabeth Schwank
Universität Münster

Kinder erwerben bereits im Vorschulalter mathematische Fähigkeiten, die für den Erfolg im Mathematikunterricht der Grundschule wesentlich sind. Dazu gehören z.B. die Kenntnis der Zahlenreihe und erster Ziffern, der Erwerb des Anzahlkonzepts und pränumerische Konzepte von Mengen und Größen (Krajewski & Ennemoser, 2018). Für Kinder mit unterdurchschnittlichen Fähigkeiten in diesen Bereichen, wäre eine gezielte Förderung bereits im Vorschulalter wünschenswert. Bei den zahlreichen erfolgreich evaluierten Förderprogrammen für den Kita-Bereich (Nafzger, 2015) haben spielbasierte Förderkonzepte (z.B. BIKO-Mathe-Kiste (Seeger et al., 2021) den Vorteil, das Spiel- und Selbstwirksamkeitsbedürfnis von Vorschulkindern aufzugreifen und sich für die Förderung den Anreiz und die Fokussierbarkeit von Regelspielen auf mathematische Inhalte zunutze machen (Hauser, 2014; Jörns et al., 2014; Seeger et al., 2018).

Ein wesentlicher Faktor für den Erfolg (spielbasierter) Förderung ist die Adaptivität der Spielleitungen (Gasteiger & Benz, 2016; Krammer, 2017). Dazu gehören adaptive Lernunterstützungsstrategien (Wood, Bruner & Ross, 1976) wie ein an das kindliche Fähigkeitsniveau adaptiertes Schwierigkeitsniveau der eingesetzten (Regel-)spiele, eine an das kindliche Verständnis adaptierte Instruktion der mathematikrelevanten Spielregeln und eine adaptive Unterstützung bei der Bewältigung mathematischer Anforderungen in den regelgeleiteten Spielzügen. Nur eine adaptive Spielbegleitung lässt Kinder Erfolge, Selbstwirksamkeit und Freude erleben, was sie zur Spielwiederholung animiert.

Aufgrund der hohen Bedeutsamkeit der Adaptivität für die Förderung mathematischer Basisfähigkeiten wurde in der vorliegenden Interventionsstudie überprüft, inwieweit sich die Adaptivität von Spielleitungen durch ein begleitendes, videobasiertes Coaching verbessern lässt. Dazu wurden die Spielleitung-Kind-Interaktion während der regelspielbasierten Matheförderung videografiert und in einstündigen Zoomsitzungen zwischen den Spieleinheiten mit der jeweiligen Spielleitung auf ihre Adaptivität hin reflektiert und bei Bedarf adaptive Handlungsalternativen entworfen.

In der vorliegenden Studie wurde ein solches videobasiertes Coaching praktiziert und mit einer strukturell gleichen Spielförderung ohne Coaching sowie einer unbehandelten Kontrollgruppe verglichen.

Die Hypothesen waren: (1) Die Intervention insgesamt fördert die mathematischen Basisfähigkeiten der Kinder im Vergleich zu einer unbehandelten Kontrollgruppe, aber in der Coachinggruppe stärker als in der Spielgruppe. (2) Spielleitungen mit Coaching verbessern ihre Adaptivität in stärkerem Maße als Spielleitungen ohne Coaching. (3) Kinder aus der Coachinggruppe verbessern ihr Regelspielverständnis und die Bewältigung mathematischer Anforderungen stärker als Kinder aus der Spielgruppe ohne Coaching.

Für die Studie wurden 119 Kinder zwischen 4 und 6 Jahren mit unterdurchschnittlichen Leistungen in mathematischen Basisfähigkeiten rekrutiert (16m und 20w in der Coachinggruppe, 17m und 17w in der Spielgruppe, 27m und 22w in der unbehandelten Kontrollgruppe). Für die Verlaufsanalysen wurden die videografierten Spielepisoden von 47 Kindern ausgewertet (29 in der Coachinggruppe, 18

in der Spielgruppe). Die regelspielbasierte Förderung folgte dem Konzept der BIKO-Mathe-Kiste (Seeger et al., 2018).

Die Förderung bestand aus neun Spieleinheiten (jeweils ca. 25 Minuten Dauer) mit drei Spielsets aus je drei Regelspielen. Spielleitungen waren 15 Psychologie- und Lehramtsstudierende, die eine Einführung zur Entwicklung mathematischer Basisfähigkeiten und zur Spielförderung erhielten. Nur die Spielleitungen der Coachinggruppe erhielten darüber hinaus zwischen den Spieleinheiten sechs Einzelcoachings.

Die summative Wirksamkeit der Förderung wurde durch ein Prä-Post-Follow-up-Design mit der Erfassung der mathematischen Basisfähigkeiten (Seeger et al., 2014) realisiert. Die Wirksamkeit des Coachings wurde durch Kodierung der videografierten Spieleinheiten anhand eines Kodiermanuals zur adaptiven Lernunterstützung ausgewertet.

Die Ergebnisse bestätigten Hypothese 1 teilweise: Kinder der Spiel- und Coachinggruppe verbesserten ihre mathematischen Basisfähigkeiten mit großem Effekt gegenüber der Kontrollgruppe, aber untereinander in vergleichbarer Weise. Hypothese 2 bestätigte sich nicht, denn die Spiel- und Coachinggruppe verbesserten ihre Adaptivität bzgl. der Instruktion in vergleichbarer Weise, während ihre Lernunterstützung gleichbleibend gut war. Hypothese 3 bestätigte sich für die Förderung des kindlichen Regelspielverständnisses, das in der Coachinggruppe signifikant stärker zunahm als in der Spielgruppe. Die Meisterung der Spiele verbesserte sich in beiden Gruppen aber vergleichbar. Die Ergebnisse werden bzgl. der Bedeutung und Wirksamkeit einer adaptiven Lernunterstützung gerade auch bei spielbasierten Förderungen diskutiert.

Keywords: adaptive Lernunterstützung, mathematische Basisfähigkeiten, Regelspiel, Programmevaluation, Kindergarten

F-12: Trends im nationalen und internationalen Large-Scale-Assessment. Schulische Bildung in bewegten Zeiten

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* S07 S00 D07

Trends im nationalen und internationalen Large-Scale-Assessment. Schulische Bildung in bewegten Zeiten

Chair(s): **Ulrich Ludewig** (Institut für Schulentwicklungsforschung, Germany), **Ramona Lorenz** (Institut für Schulentwicklungsforschung, Germany), **Nele McElvany** (Institut für Schulentwicklungsforschung, Germany)
*Diskutant*in(nen):* **Frank Goldhammer** (Leibniz-Institut für Bildungsforschung und Bildungsinformation)

Das Symposium „Trends im nationalen und internationalen Large-Scale-Assessment. Schulische Bildung in bewegten Zeiten“ befasst sich mit Ergebnissen aus Large-Scale-Assessments (LSA) zu aktuellen gesellschaftlichen Herausforderungen und zentralen Bildungsoutcomes. Berichterstattung über LSAs haben in den vergangen zwei Jahrzehnten die öffentlichen Debatten zu bildungsrelevanten Themen entscheidend geprägt. Zwei wichtige aktuelle Herausforderungen sind zum einen die Folgen der corona-bedingten Unterrichtseinschränkungen und zum anderen der Umgang mit dem Klimawandel. Die vier Beiträge des Symposiums berichten aus vier wichtigen großangelegten Studien in Deutschland, dem Bildungstrend 2021, der internationalen Grundschul-Lese-Untersuchung (IGLU) 2021, dem Programme for International Student Assessment (PISA) 2018 und dem Nationalen Bildungspanel (NEPS). Im Symposium werden verschiedene Bildungsoutcomes und Rahmenbedingungen mit großen zu meist repräsentativen Datensätzen untersucht, die dafür genutzt werden Trends zu identifizieren und zu erklären oder um detaillierte Erklärungsmodelle für bildungsrelevante Prozesse zu entwickeln.

Schneider et al. berichten aus dem IQB-Bildungstrend zur Lese-, Zuhör-, Orthographie- und Mathematikkompetenz aus repräsentativen Erhebungen in den Jahren 2011, 2016 und 2021 bei Viertklässler:innen. Die aktuellen Ergebnisse aus dem Jahr 2021 werden vor dem Hintergrund der vorherigen Studien interpretiert, mit einem Fokus auf Bildungsstandards, Subgruppenunterschiede und den Zusammenhang zwischen Kompetenzen und Unterrichtseinschränkungen.

Ludewig et al. schließen an das Thema an und berichten zum langfristigen Trend der Lesekompetenz bei Viertklässler:innen aus IGLU 2001 bis 2021. Im Fokus des Beitrags steht die Frage, welche Rolle die Schüler:innenkomposition bei der Erklärung des Trends spielt. Berücksichtigung finden z.B. die Entwicklung des Anteils von Kindern mit Zuwanderungshintergrund oder das Bildungsniveau der Eltern. Des Weiteren wird abgeschätzt, in wie weit sich die Ergebnisse nach corona-bedingten Unterrichtseinschränkungen im Jahr 2021 vom langfristigen Trend abgrenzen lassen.

Nusser et al.‘s Beitrag befasst sich mit der Entwicklung mathematischer Kompetenzen bei Sekundarschüler:innen anhand zweier Kohorten, von denen eine durch corona-bedingte Unterrichtseinschränkungen betroffen war. Beide Kohorten nahmen in der 7. und 9. Klasse an einem Mathematiktest teil. Im Fokus steht die Abschätzung der Folgen von corona-bedingten Unterrichtseinschränkungen und das Ausmaß, mit dem motivationale und kognitive Merkmale einen möglichen Effekt moderieren.

Diedrich et al. befassen sich mit dem Klimabewusstsein Fünfzehnjähriger in Deutschland im internationalen Vergleich auf der Basis von PISA 2018. In Deutschland liegt die umweltbezogene Informiertheit mit 82.6 Prozent über dem Durchschnitt der OECD-Staaten. Die umweltbezogenen

Aktionen liegen in Deutschland jedoch unter dem OECD-Durchschnitt. In einer vertiefenden Analyse werden Zusammenhänge zwischen den umweltbezogenen Aktionen der Jugendlichen mit ihrer umweltbezogenen Informiertheit, ihren Werthaltungen sowie Unterrichts- und Hintergrundmerkmalen untersucht.

Beiträge des Symposiums

Ergebnisse des IQB-Bildungstrends 2021: Hängen sprachliche und mathematische Kompetenzen von Viertklässler:innen mit den Lernbedingungen während der Coronavirus-Pandemie zusammen?

Rebecca Schneider, Stefan Schipolowski, Karoline A. Sachse, Florian Enke, Petra Stanat
Institut zur Qualitätsentwicklung im Bildungswesen

[HINWEIS: Aufgrund einer Sperrfrist bis zum 17.10.2022 können im Abstract keine Ergebnisse genannt werden.]

Mit den IQB-Bildungstrends wird regelmäßig überprüft, inwieweit Schüler:innen in Deutschland die mit den Bildungsstandards der Kultusministerkonferenz definierten Kompetenzziele erreichen. Im Primarbereich werden diese Studien am Ende der 4. Jahrgangsstufe in den Fächern Deutsch und Mathematik in allen 16 Ländern in der Bundesrepublik Deutschland durchgeführt. Nach der ersten Studie im Jahr 2011 (IQB-Ländervergleich 2011; Stanat et al., 2012) und der zweiten Erhebung im Jahr 2016 (IQB-Bildungstrend 2016; Stanat et al., 2017) wurden die Kompetenzen der Viertklässler:innen im Jahr 2021 (IQB-Bildungstrend 2021; Stanat et al., 2022) bereits zum dritten Mal untersucht. An der Erhebung im Jahr 2021 nahm eine auf Bundes- und Länderebene repräsentative Stichprobe von insgesamt 26.844 Schüler:innen aus 1.464 Schulen teil. Wie bereits in den Vorerhebungen wurden neben den Kompetenztests auch umfangreiche Befragungen der teilnehmenden Schüler:innen, ihrer Eltern, Lehrkräfte und Schulleitungen durchgeführt.

Die Erhebungen zum IQB-Bildungstrend 2021 fanden unter außergewöhnlichen Bedingungen statt. Zu Beginn der Pandemie und der ersten Schulschließungen im Frühjahr 2020 besuchten die am Bildungstrend teilnehmenden Kinder die 3. Jahrgangsstufe, zum Zeitpunkt der Datenerhebung zwischen April und August 2021 konnten viele von ihnen nach dem ersten Pandemiewinter und verschiedenen Phasen von Präsenz- und Fernunterricht, Teilungsunterricht und Unterricht im vollständigen Klassenverband erst seit wenigen Wochen wieder regelmäßig ihre Schule besuchen. Um dieser besonderen Situation Rechnung zu tragen, wurden in den Befragungen der Schulleitungen, Lehrkräfte und Eltern verschiedene Aspekte des Fern- bzw. Wechselunterrichts und der Lernbedingungen zu Hause erfasst. So wurden die Schulleitungen um Angaben zum Umfang von Fern- und Wechselunterricht zwischen März 2020 und dem schulspezifischen Testtermin gebeten. Die Eltern sollten Angaben zur räumlichen und technischen Ausstattung (z. B. ein Raum zum ungestörten Lernen, für den Fernunterricht ausreichendes Internet) der Kinder zu Hause machen und einschätzen, inwieweit sie ihre Kinder beim häuslichen Lernen unterstützen konnten. Zudem sollten Eltern und Lehrkräfte einschätzen, wie gut verschiedene Aspekte des Fernunterrichts funktioniert haben (z. B. Übermittlung von Arbeitsaufträgen und Lernmaterialien).

In diesem Beitrag soll zunächst überblicksartig zusammengefasst werden, wie sich die von Viertklässler:innen in Deutschland insgesamt und in den Ländern erreichten Kompetenzen zwischen

den Jahren 2011, 2016 und 2021 in den Fächern Deutsch (Lesen, Zuhören, Orthografie) und Mathematik (Globalskala) verändert haben. Neben Trends im Erreichen von Mindest- und Regelstandards in den verschiedenen Kompetenzbereichen werden auch Veränderungen des Kompetenzniveaus in verschiedenen Teilpopulationen untersucht, wobei Kinder aus sozial weniger gut gestellten Familien und Kinder aus zugewanderten Familien im Vordergrund stehen. Der zweite Teil des Beitrags informiert über die Befragungsergebnisse zum Fern- und Wechselunterricht und zu den Lernbedingungen der Schüler:innen während der Coronavirus-Pandemie. Zudem wird mit Hilfe von Mehrebenen-Regressionsanalysen der Frage nachgegangen, inwieweit die erreichten Kompetenzen der Schüler:innen im Jahr 2021 mit den erfassten Merkmalen des Fern- bzw. Wechselunterrichts und den Lernbedingungen während der Pandemie zusammenhängen.

In der abschließenden Diskussion erfolgt eine Einordnung der Ergebnisse vor dem Hintergrund anderer nationaler und internationaler Befunde zu den Auswirkungen der Pandemie auf die Lernentwicklung von Kindern. Zudem werden Implikationen der Befunde für die Akteure im Bildungssystem diskutiert.

Welche Rolle spielt die Schüler:innenkomposition für den IGLU Lesekompetenztrend zwischen 2001 und 2021?

Ulrich Ludewig, Ramona Lorenz, Ruben Kleinkorres, Rahim Schaufelberger, Nele McElvany
Institut für Schulentwicklungsforschung, Germany

Die Beschreibung langfristiger Entwicklungen von Schulleistungen gehört mitunter zu den bildungspolitisch und gesellschaftlich relevantesten Aufgaben der empirischen Bildungsforschung (Klieme, 2020). Die internationale Grundschul-Lese-Untersuchung (IGLU) ist eine der ersten der großen Large-Scale-Assessments (LSA) in Deutschland, die von ihrem ersten bis zu ihrem aktuellsten Erhebungszyklus einen Gesamtzeitraum von 20 Jahren umfasst. Zu Beginn der Nuller-Jahre verzeichnete sie einen ansteigenden, gefolgt von einem nach 2006 sinkenden oder stagnierenden Trend bei der Lesekompetenz von Viertklässler:innen in Deutschland (Hußmann et al., 2017). Eine entscheidende Frage für die Interpretation des Trends über 20 Jahre und insbesondere für die Einordnung der Ergebnisse im Jahr 2021, nach eingeschränktem Unterricht während der COVID-19 Pandemie (König & Frey, 2022; Ludewig et al., 2022; Stanat et al., 2022): Welche Rolle spielte die Schüler:innenkomposition für die Trendergebnisse. Die Lesekompetenz ist, wie andere Domänen (z.B. Reis et al., 2019), wesentlich mit dem familiären Hintergrund und den damit verbundenen Lernvoraussetzungen assoziiert. Wenn sich die Trendergebnisse durch veränderte familiäre Hintergrundmerkmale einer Schüler:innenkohorte erklären ließen, würde dies nahelegen, dass die Entwicklung der Lesekompetenz eher mit veränderten Lernvoraussetzungen als mit einer veränderten Qualität im Bildungssystem zusammenhängt. In den letzten 20 Jahren gab es bedeutende demographische Entwicklungen (Bildungsberichterstattung, 2020), die sich positiv und negativ auf die durchschnittliche Lesekompetenz der Viertklässler:innen in Deutschland ausgewirkt haben könnten. Einen eher positiven Einfluss könnten höhere Bildungsabschlüsse und bessere Beschäftigungsverhältnisse der Eltern gehabt haben. Steigende Anteile von Kindern mit Zuwanderungshintergrund oder geringere Präsenz von gedruckten Büchern in den Haushalten könnten einen eher negativen Einfluss haben. In dem Beitrag wird unter Einbezug verschiedener Modellrechnungen eingeordnet (i) wie stark familiäre Hintergrundmerkmale den Trend über 20 Jahre

beeinflusst haben könnten und (ii) wie der Lesekompetenztrend mit konstanten familiären Hintergrundmerkmalen aussehen würde.

Grundlage des Beitrags sind die Daten Deutschlands von IGLU 2001 (Bos et al., 2002), 2006 (Bos et al., 2006), 2011 (Bos et al., 2012), 2016 (Hußmann et al., 2017) und 2021 (McElvany et al., in prep) mit insgesamt N = 28 102 Viertklässer:innen (Alter: M = 10.4 (0.5) Jahre, Anteil Mädchen: 50 %). Die Aufbereitung der Daten aus den fünf Studien erfolgte mit Aufmerksamkeit auf die Vergleichbarkeit der gewährten Indikatoren (z.B. hinsichtlich geänderter Formulierungen). Die Trends der familiären Hintergrundmerkmale wurden mit Hilfe von Ergebnissen aus anderen repräsentativen Studien (z. B. IQB Bildungstrend, Mikrozensus) auf Plausibilität geprüft. Um den Einfluss von Nicht-Antworten auf die Ergebnisse möglichst zu verringern wurden multiple imputierte Datensätze mit Hilfe der R-Paketes miceadds (Robitzsch et al., 2017) erstellt. Für das Imputationsmodell wurden Hilfsvariablen aus Schüler:innen-, Eltern- und Schulleitungsfragebögen verwendet und die Gruppierung in Schulklassen berücksichtigt. Die Analyse erfolgte unter Einbezug der multiple imputierten Daten, pausibler Werte der Lesekompetenz und der Schüler:innengewichtung anhand eines Mixed-Effect Modelles für komplexe Survey Daten mit Hilfe das R-Paketes BIFIE.survey (Robitzsch & Oberwimmer, 2022).

Der Ergebnisbericht unterliegt bis zur Veröffentlich des IGLU-2021-Berichts am 13.12.2022 einer Sperrfrist und kanndaher an dieser Stelle noch nicht beschrieben werden. Die Ergebnisse werden nach dieser Sperrfrist im Abstract ergänzt.

Eine Limitation der Studie ist, dass anhand der querschnittlichen Daten kein direkter Rückschluss auf die Veränderung der Qualität des Bildungssystems möglich ist. Des Weiteren lassen sich nicht alle wichtigen Trends effektiv mit den Daten abbilden (z.B. bundeslandspezifische Reformeffekte, Ausbau von Kindertagestätten). Zu den Stärken des Beitrags gehören das relativ große Set an familiären Hintergrundmerkmalen, die für die Modellierung herangezogen wurden, sowie die Anwendung elaborierter Analysemethoden, die die komplexe Datenstruktur angemessen berücksichtigen.

Der Beitrag demonstriert, dass Trends bei familiären Hintergrundmerkmalen wichtige Aspekte bei der Interpretation von Lesekompetenztrendergebnissen in LSAs sind, die bei der Berichterstattung häufig nicht systematisch berücksichtigt werden. Zur vertiefenden Analyse streben die Autoren:innen eine Replikation der Ergebnisse mit weiteren repräsentativen LSA Studien an.

Mathematische Kompetenzentwicklung von Schüler:innen in der Sekundarstufe während der corona-bedingten Schulschließungen – ein NEPS-Kohortenvergleich

Lena Nusser, Kathrin Lockl, Timo Gnambs, Ilka Wolter, Cordula Artelt
Leibniz-Institut für Bildungsverläufe

Die Maßnahmen zur Eindämmung der Corona-Pandemie beinhalteten massive Eingriffe in das Schulsystem. Schließungen der Bildungseinrichtungen und der damit einhergehende Distanzunterricht führten wiederholt zu Veränderungen der individuellen Lernsituation der Schüler:innen. Damit verbundene Herausforderungen für alle beteiligten Akteur:innen wie Schulleitungen, Lehrkräfte und Schülerschaft sowie ihre Eltern wurden vielfach diskutiert und dokumentiert (Fickermann & Edelstein, 2021; Gehrer et al., 2022; Huber et al., 2021). Ergebnisse aus der Zeit der ersten bundesweiten Schulschließungen im Frühjahr 2020 zur Altersgruppe der Fünfzehnjährigen deuten darauf hin, dass Schüler:innen insgesamt deutlich weniger Zeit in das

Lernen investierten im Vergleich zur regulären Schulzeit (Nusser et al., 2021). Dies traf insbesondere auf Schüler:innen zu, die vorher geringere Kompetenzen aufwiesen als ihre Altersgenoss:innen (Nusser, 2021). Eltern, die in dieser Zeit befragt wurden, nahmen an, dass nur etwa ein Drittel der Kinder während der Schulschließungen vergleichbar oder mehr in den Hauptfächern gelernt hat als zu regulären Schulzeiten (Nusser et al., 2021). Auch wenn laut Elternberichten die Schüler:innen insgesamt recht gut mit der neuen Situation des Lernens zuhause umgehen konnten, so war dies positiv mit der vorherigen Lesekompetenz und dem Lernverhalten wie Anstrengungsbereitschaft assoziiert (Lockl et al., 2021). Trendanalysen vom IQB und IGLU legen nahe, dass Grundschulkinder gravierende Lernrückstände von etwa einem Drittel Schuljahr aufweisen und mehr Schüler:innen als früher die definierten Mindeststandards nicht erreichten (Ludwig et al., 2022, Stanat et al., 2022). Auch weitere (inter)nationale Studien zeigen, dass die Kompetenzentwicklung der Schüler:innen durch die Schulschließungen negativ beeinflusst wurde (Moliner & Alegre, 2022; Panagouli et al., 2021; Schult et al., 2022).

Ob ähnliche negative Effekte auch für ältere Schüler:innen am Ende der Pflichtschulzeit zu beobachten sind, ist noch nicht hinreichend untersucht. Dafür, dass die Effekte in dieser Altersgruppe geringer ausfallen könnten, spricht z.B., dass die Fähigkeiten zum selbstregulierten Lernen in der Sekundarstufe stärker ausgeprägt sind im Vergleich zu Grundschulzeit (Pintrich & Zusho, 2002); gleiches gilt vermutlich für den Umgang mit digitalen Medien und deren Nutzung zum Lernen, was den Zugang zum Distanzunterricht möglicherweise erleichterte. Die vorliegende Studie untersucht in einem Kohortenvergleich, inwieweit die deutschlandweiten Schulschließungen die mathematische Kompetenzentwicklung von Schüler:innen zwischen der 7. und 9. Klasse beeinflusst haben. Darüber hinaus wird geprüft, ob und in welchem Ausmaß sowohl motivationale als auch kognitive Merkmale einen möglichen Effekt moderieren.

Für die Analysen werden jeweils zwei Wellen der Startkohorten 2 (Kindergarten) und 3 (Sekundarstufe I) des Nationalen Bildungspanels (NEPS; Blossfeld & Roßbach, 2019) herangezogen, in denen die mathematische Kompetenz in den Klassenstufen 7 und 9 erfasst wurde. Die interessierenden Messungen in Startkohorte 2 erfolgten im ersten Halbjahr des Schuljahrs 2018/19 (Klassenstufe 7), also vor den corona-bedingten Schulschließungen, sowie im Zeitraum von Juni bis Oktober 2021 (Klassenstufe 9) während der Corona-Pandemie. Bei der Vergleichsstichprobe der Startkohorte 3, welche keine Phasen systematischer Schulschließungen erlebte, fanden die Erhebungen im Schuljahr 2012/2013 (Jahrgangsstufe 7) und zwei Jahre später im Schuljahr 2014/2015 (Jahrgangsstufe 9) statt. Somit erlebte die jüngere Kohorte erschwerte Lernbedingungen aufgrund der corona-bedingten Schulschließungen, während für die ältere Kohorte eine Kompetenzentwicklung unter regulären Bedingungen zu beobachten war. Die beiden Stichproben wurden anhand relevanter Hintergrundmerkmale (z.B. Kompetenz, sozioökonomischer Status) statistisch gematcht, um bei den Analysen auf vergleichbare Ausgangsstichproben in der siebten Klassenstufe zurückgreifen zu können. Dies resultierte in einer Analysestichprobe von ca. N = 4000 Sekundarschüler*innen. Die Erstveröffentlichung der Ergebnisse erfolgt im Herbst 2022. Im Rahmen der Konferenz werden die Ergebnisse zum Einfluss der corona-bedingten Schulschließungen auf die Kompetenzentwicklung präsentiert sowie ihre Implikationen für die Lernsituation an Schulen diskutiert.

Klimabewusstsein Fünfzehnjähriger in Deutschland: vom Wissen und Können übers Wollen zum Tun

Jennifer Diedrich[1], Julia Mang[1], Sabine Patzl[1], Sophie Sessler[1], Monika Martin[2], Doris Lewalter[1]
[1]Zentrum für internationale Bildungsvergleichsstudien (ZIB), Technische Universität München, Deutschland, [2]Pädagogische Hochschule Freiburg, Deutschland

Theoretischer Hintergrund:

Der menschengemachte Klimawandel gehört zu den drängendsten Problemen der heutigen Zeit und bedarf bewusster Handlungen auf allen Ebenen von staatlich bis privat (UNFCCC, 2015). Wie alle Handlungen haben auch umweltbezogene Aktionen Voraussetzungen. Handlungstheorien (z.B. Ajzen, 1991) beschreiben als Vorbedingung dafür, dass eine Person aktiv wird, dass sie genug weiß (Wissen) und zu der Handlung motiviert ist (Wollen). In PISA 2018 wurden diese Aspekte operationalisiert über umweltbezogene Aktionen (Tun), umweltbezogene Informiertheit (Wissen) sowie umweltbezogene Werthaltungen (Wollen). Diese werden unter dem Begriff Klimabewusstsein subsummiert.

Fragestellung:

Die Studie untersucht das Klimabewusstsein Jugendlicher in Deutschland mit einem Fokus auf den umweltbezogenen Aktionen. Zunächst soll untersucht werden, welche Ausprägung die Merkmale des Klimabewusstseins im Vergleich zu ausgewählten Staaten sowie dem OECD-Durchschnitt haben. Anschließend wird der Frage nachgegangen, ob sich ausgewählte Gruppen innerhalb Deutschlands (Mädchen vs. Jungen; Schularten) in ihrer umweltbezogenen Informiertheit, Werthaltungen sowie Aktionen unterscheiden.

Methode:

Dazu werden PISA Daten einer Sekundäranalyse unterzogen. Für PISA 2018 wurden 5451 repräsentativ gezogene Fünfzehnjährige (50 % Mädchen; 59.7 % nicht-gymnasiale Schularten) zu den oben genannten Merkmalen befragt. Zusätzlich beantworteten ihre Eltern, Schulleitungen sowie Lehrkräfte ihrer Schule Kontextfragebögen. Eine genaue Stichprobenbeschreibung findet sich in der Primärstudie (Mang et al., 2019; Weis & Reiss, 2019). Zur umweltbezogenen Informiertheit wurden die Jugendlichen auf einer vierstufigen Skala gefragt, wie gut sie sich über „Klimawandel und globale Erwärmung“ informiert fühlen. Der Aussage „Die globale Umwelt zu schützen ist mir wichtig“ stimmten Jugendliche auf einer vierstufigen Skala zu, die über umweltbezogene Werthaltungen verfügen. Umweltbezogene Aktionen wurden mittels fünf Items dichotom erfasst, u.a. „ich reduziere den Energieverbrauch zuhause“ und „ich beteilige mich an Aktivitäten zugunsten des Umweltschutzes“.

Ergebnisse und ihre Bedeutung:

Die umweltbezogene Informiertheit liegt mit 82,6 Prozent in Deutschland über dem Durchschnitt der OECD-Staaten (78.9 %). Diese ist bei den Jungen stärker ausgeprägt als bei Mädchen und bei Schüler:innen an Gymnasien stärker als an nicht-gymnasialen Schularten. Umweltbezogene Werthaltungen wiederum werden von Fünfzehnjährigen in Deutschland (67.4 %) seltener berichtet als im Durchschnitt der OECD-Staaten (77.9 %). Auch das Wollen wird an Gymnasien häufiger berichtet als an anderen Schularten. Niedriger sind die umweltbezogenen Aktionen ausgeprägt, diese werden alle in Deutschland seltener als im OECD-Durchschnitt berichtet. In einem nächsten Analyseschritt,

welcher bis zur Konferenz abgeschlossen sein wird, werden Kontextvariablen hinsichtlich ihrer Zusammenhänge mit den Aktionen der Jugendlichen verglichen. Dazu werden die Aktionen der Jugendlichen mittels Odds-Ratios durch ihre umweltbezogene Informiertheit und Werthaltungen, Geschlecht, Migrationshintergrund, eigene umweltbezogene Aktionen der Eltern sowie das Vorkommen der Klimawandels in Lehrplan und Unterricht vorhergesagt. Im Rahmen des Vortrags sollen Ansätze diskutiert werden, wie Jugendliche vom Wissen übers Wollen ins Tun kommen können.

F-13: Potenziale geskripteter Unterrichtsvideos für die Lehrkräftebildung

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* T03 R02 D39

Potenziale geskripteter Unterrichtsvideos für die Lehrkräftebildung

Chair(s): **Julia Bönte** (Universität Duisburg-Essen), **Tosca Daltoè** (Universität Tübingen & Institut für Bildungsanalysen Baden-Württemberg, Stuttgart)
*Diskutant*in(nen):* **Marc Kleinknecht** (Leuphana Universität Lüneburg)

Während die universitäre Phase der Lehrkräftebildung lange Zeit die Förderung des Fachwissens sowie des fachdidaktischen Wissens fokussierte, wurde die Vorbereitung auf die Berufspraxis bislang weniger adressiert (Kolthoff & Terhart, 2013). Um der mangelnden Verzahnung von Theorie und Praxis entgegenzuwirken und angehende Lehrkräfte frühzeitig praxisorientiert auf den Berufsalltag vorzubereiten, werden in der Lehrkräfte(aus)bildung heute verstärkt videogestützte Analyse- und Reflexionsformate genutzt (Borko et al., 2011; Seidel & Thiel, 2017). Speziell in der Lehrkräfteausbildung können Unterrichtsvideos zu einer praxisnahen Vorbereitung auf den Lehrberuf beitragen, indem reale Fälle aus der Unterrichtspraxis abgebildet werden können, ohne Lehramtsstudierende dem Handlungsdruck der tatsächlichen Unterrichtssituation auszusetzen (Gold, Förster & Holodynski, 2013; van Es und Sherin, 2008). Unterrichtsvideos ermöglichen beispielsweise die Schulung der professionellen Unterrichtswahrnehmung im Unterrichtskontext (Erkennen, wissensbasierte Interpretation und Handlungsalternativengenerierung; z. B. Gold et al., 2017; Steffensky, Gold, Holodynski & Möller, 2015) oder der Klassenführungsexpertise (Kramer et al., 2017).

Neben Videoaufzeichnungen realen Unterrichts werden in der Lehrkräftebildung zunehmend auch Videos geskripteter Unterrichtssituationen eingesetzt. Geskriptete bzw. inszenierte oder auch staged Videos bieten dabei besondere Vorteile zur Repräsentation und Analyse sowie Reflexion von Unterrichtssituationen, z. B. durch eine Fokussierung auf ausgewählte Aspekte des Unterrichtsgeschehens oder bestimmte konzeptionelle Schwerpunkte, aber auch die Repräsentation von sensiblen oder sehr kritischen Situationen (Gartmeier, 2014). Trotz Inszenierung kann dabei mittels gezielter Entwicklungsarbeit eine hohe empfundene Authentizität erzielt werden (z. B. Kramer et al., 2020; van Bebber, 2021).

Das vorliegende Symposium fokussiert die Entwicklung und Evaluation geskripteter Unterrichtsvideos in Hinblick auf ihr Potenzial für die Lehrkräftebildung. Der erste Beitrag stellt exemplarisch die Entwicklung geskripteten kontrastierenden Videomaterials dar, wobei im Rahmen einer Begleiterhebung zur Videoproduktion die am Videodreh beteiligten Klassen die geskripteten Unterrichtssituationen hinsichtlich Authentizität und Unterrichtsqualität einschätzen.

Im zweiten Beitrag wird eine Evaluationsstudie präsentiert, in der die zuvor entwickelten geskripteten CLIPSS-Videovignetten zum Schwerpunkt Klassenführung in Hinblick auf Sensitivität und Spezifität bezüglich des Erkennens von (nonverbalen) Klassenführungsstrategien bei der Videobetrachtung untersucht werden. Ergänzend präsentiert werden Gruppenvergleiche zwischen Noviz*innen und praktizierenden Lehrkräften hinsichtlich Sensitivität und Spezifität sowie Authentizität, Immersion und Resonanz.

Im Rahmen einer Vergleichsstudie untersucht der dritte Beitrag schließlich die Effektivität instruktionaler Unterstützungsansätze zur Förderung der professionellen Unterrichtswahrnehmung

mittels geskripteter Videos. Dabei werden unterstützende Videoannotationen zur Qualitätsförderung der Unterrichtsreflexion mit dem thematischen Schwerpunkt der motivationalen Aktivierung eingesetzt.

Das Symposium liefert weiterführende Erkenntnisse zur Entwicklung, zur Wahrnehmung und zu den Einsatzmöglichkeiten geskripteter Unterrichtsvideos. Übergeordnet soll diskutiert werden, welche Potenziale geskriptete Unterrichtsvideos für die Lehrkräftebildung bergen.

Beiträge des Symposiums

Geskriptete Unterrichtsvideos: Wahrnehmung von Authentizität und Unterrichtsqualität des inszenierten Unterrichts bei der Videoproduktion aus Sicht der Beteiligten

Jana Philipp[1], Tosca Daltoè[2], Julia Larissa Maier[1], Evelin Ruth-Herbein[1], Benjamin Fauth[2]
[1]Institut für Bildungsanalysen Baden-Württemberg, Stuttgart, [2]Universität Tübingen & Institut für Bildungsanalysen Baden-Württemberg, Stuttgart

Theoretischer Hintergrund

Der Einsatz geskripteter Unterrichtsvideos in Lehrkräftebildung und Unterrichtsforschung ist im Vergleich zum Einsatz realer Unterrichtsvideos bislang eher die Ausnahme. Inszenierte Unterrichtsvideos haben im Vergleich zu realen Unterrichtsvideos den Vorteil, dass durch vorgegebene Handlungsabläufe bestimmte Unterrichtsmerkmale gezielt fokussiert werden können (Kim et al., 2006). Die Lehrkräftebildung würde von Unterrichtsvideos zur Verdeutlichung von qualitätsvollen und weniger qualitätsvollen Unterrichtssituationen profitieren (Seidel et al., 2021). Im Videoprojekt zum Unterrichtsfeedbackbogen Tiefenstrukturen (UFB; Fauth et al., 2021) werden inszenierte Unterrichtsvideos erstellt, die die zentralen Unterrichtsqualitätsmerkmale des UFB mit Positiv- und Negativbeispielen gezielt veranschaulichen (vgl. Seidel et al., 2021). Der UFB wurde als wissenschaftlich fundiertes Unterrichtsbeobachtungsinstrument für das Bildungssystem in Baden-Württemberg entwickelt (Fauth et al., 2021; Maier et al., 2021). Anhand elf hoch-inferenter Items erfasst der UFB zentrale Merkmale der Tiefenstrukturen des Unterrichts (kognitive Aktivierung, konstruktive Unterstützung, strukturierte Klassenführung). Einsatzbereiche des UFB sind kollegiales Unterrichtsfeedback, Selbstreflexion sowie Lehrkräftebildung. Die Unterrichtsvideos sollen den erfolgreichen Einsatz des UFB unterstützen.

Als wichtiger Qualitätsaspekt von Unterrichtsvideos gilt die empfundene Authentizität (Gartmeier, 2014). Bisherige Forschung hat gezeigt, dass die Inszenierung von Unterricht in Videos nicht zwingend zu einem Mangel an empfundener Authentizität führt (Kramer et al., 2020; Piwowar et al., 2018). Wissenschaftliche Untersuchungen von inszenierten Unterrichtsvideos fokussieren bislang die Wahrnehmung der Videos durch Lehramtsstudierende (Codreanu et al., 2020; Piwowar et al., 2018). Die Wahrnehmung der inszenierten Unterrichtssituationen der an der Videoproduktion teilnehmenden Klassen ist hingegen bisher unerforscht. Die empfundene Authentizität der teilnehmenden Klassen sowie deren Wahrnehmung der Unterrichtsqualität könnte jedoch Hinweise zur Qualität der Inszenierung der Unterrichtssituationen geben.

Fragestellung

Im vorliegenden Beitrag wird untersucht, (1) wie am Videodreh teilnehmende Klassen inszenierten Unterricht hinsichtlich Authentizität wahrnehmen; (2) wie die Unterrichtsqualität inszenierten

Unterrichts aus Perspektive der teilnehmenden Schülerinnen und Schüler eingeschätzt wird; sowie (3) inwieweit gezielte Manipulationen in der Unterrichtsqualität der inszenierten Unterrichtssituation von den teilnehmenden Schülerinnen und Schülern und der Lehrkraft erkannt werden.

Methode

Die Videoproduktion orientiert sich am schrittweisen Ansatz zur Produktion inszenierter Unterrichtsszenarien nach Dieker et al. (2009). Angelehnt an Dieker et al. (2009) werden zur Videoproduktion drehbuchartige Unterrichtsskripte erstellt. Die Skripte beschreiben jeweils eine Unterrichtssituation (ca. 10 Min.) und verdeutlichen verschiedene Ausprägungen der Unterrichtsqualitätsmerkmale des UFB. Dabei werden unterschiedliche Versionen derselben Unterrichtssituation erstellt, die auf einzelnen Unterrichtsqualitätsmerkmalen manipuliert sind, während die anderen Qualitätsmerkmale konstant bleiben. Die inszenierten Unterrichtsvideos werden dann, auf Grundlage der Skripte, mit Schulklassen aufgezeichnet. Durch die enge Einbindung von erfahrenen Lehrkräften in den Prozess der Videoproduktion sollen möglichst authentische und alltagsnahe Unterrichtsvideos entstehen (vgl. Kramer et al., 2020).

Durch Fragebogenerhebungen nach der Videoaufzeichnung werden die am Dreh beteiligten Schülerinnen und Schüler sowie die Lehrkraft zur Unterrichtsqualität der gefilmten Unterrichtssituation (UFB) und zur Authentizität der Unterrichtssituation befragt. Ferner werden Angaben zu Bedingungen der Aufzeichnung der inszenierten Unterrichtsvideos sowie personenbezogene Daten erhoben.

Erwartete Ergebnisse und ihre Bedeutung

Im Vortrag werden Ergebnisse zur Wahrnehmung der Unterrichtssituationen durch die Schülerinnen und Schüler hinsichtlich Authentizität (vier Subskalen, α = .657-.914) und Unterrichtsqualität (UFB; Fauth et al., 2021) präsentiert. Dabei werden neben den bereits im Schuljahr 2022/2023 erhobenen Daten (drei Unterrichtsvideos, jeweils N = 24 Schülerinnen und Schüler) weitere Daten aus im aktuellen Schulhalbjahr stattfindenden Videoaufzeichnungen und Begleiterhebungen (fünf Unterrichtsvideos) einfließen.

Die Ergebnisse aus dem Videoprojekt zum UFB können zur Qualitätskontrolle der inszenierten Unterrichtsvideos beitragen. Darüber hinaus können Hinweise zu Gelingensbedingungen der Produktion inszenierter Videos abgeleitet werden. Die inszenierten Unterrichtsvideos aus dem Videoprojekt zum UFB können Maßnahmen zur Unterrichtsqualitätsentwicklung in der Lehrkräfteaus- und -fortbildung unterstützen.

Keywords: Unterrichtsqualität, Unterrichtsvideos, geskriptete Videos, Authentizität

Evaluation geskripteter Unterrichtsvideovignetten hinsichtlich Sensitivität und Spezifität bezüglich des Erkennens von (nonverbaler) Klassenführung

Julia Bönte[1], Gerlinde Lenske[2], Detlev Leutner[1]
[1]Universität Duisburg-Essen, [2]Leuphana Universität Lüneburg

Theoretischer Hintergrund

Die Klassenführung bildet eine zentrale Komponente für die Realisierung gelungener Unterrichtsprozesse (z. B. König & Pflanzl, 2016). Sie kann dabei fächer- und schulformübergreifend positiven Einfluss auf das Unterrichtsgeschehen nehmen. Es existieren u. a. positive

Zusammenhänge der Klassenführung mit dem Wohlbefinden von Lehrkräften (Dicke et al., 2015), der Schülerleistung (Lenske et al., 2016) und dem situationalen Interesse der Lernenden (Lenske, Wirth & Leutner, 2017). Insbesondere angehende sowie unerfahrene Lehrkräfte, sog. Noviz*innen, fühlen sich speziell auf den Anforderungsbereich Klassenführung unzureichend vorbereitet (Evertson & Weinstein, 2006).

Videogestützte Trainingsformate bieten neben der Schulung der professionellen Wahrnehmung (Steffensky, Gold, Holodynski & Möller, 2015) auch Potenziale zur Förderung der Klassenführungskompetenz (z. B. Hellermann, Gold & Holodynski, 2015). Bezüglich der Klassenführungskompetenz bestehen Unterschiede zwischen Expert*innen und Noviz*innen (Thiel, Richter & Ophardt, 2012). Die visuelle Aufmerksamkeit von Expert*innen zur Identifikation von relevanten Informationen im Unterrichtsgeschehen ist zudem fokussiert und wissensgesteuert, wohingegen die visuelle Aufmerksamkeit von Noviz*innen wenig fokussiert und eher diffus ist (Wolff et al., 2016). Expert*innen sind Noviz*innen bei der Analyse komplexer Unterrichtssituationen daher überlegen.

Im Projekt CLIPSS – CLassroom management in Primary and Secondary Schools (Teilprojekt der Qualitätsoffensive Lehrerbildung – QLB; BMBF, 2016) wurden inszenierte Unterrichtsvideovignetten zum Schwerpunkt Klassenführung mit besonderem Fokus auf nonverbale Klassenführungskomponenten entwickelt, um ein Erkennen klassenführungsrelevanter Unterrichtsaspekte sowohl für Noviz*innen als auch Expert*innen zu ermöglichen. Durch die gezielte Drehbuchentwicklung und inszenierte, besonders fokussierte, Darstellung sollte es Lehramtsstudierenden ebenso wie praktizierenden Lehrkräften ermöglicht werden, die in den Videovignetten intendierten Klassenführungsstrategien auf Basis des Linzer Konzepts der Klassenführung (LKK; Lenske & Mayr, 2015) und nonverbalen Klassenführungskomponenten (z. B. Rosenbusch, 2004) korrekt zu erkennen. Dabei standen die folgenden Forschungsfragen im Fokus: (1) Gelingt das korrekte Erkennen von Klassenführungsstrategien und nonverbalen Klassenführungskomponenten als in den CLIPSS-Videovignetten fokussiert (Sensitivität) und als nicht in den Videovignetten fokussiert (Spezifität)? (2) Unterscheiden sich die Sensitivität und die Spezifität zwischen Noviz*innen und praktizierenden Lehrkräften?

Methode

Zur Evaluation des Videomaterials schätzten die Proband*innen (N = 243 Bachelorlehramtsstudierende, MAlter = 21.27 Jahre, SD = 3.15, MSemester = 2.60, SD = 1.85, 68 % weiblich; N = 18 praktizierende Lehrkräfte, MAlter = 42.72 Jahre, SD = 11.49, MLehrerfahrung = 17.92 Jahre, SD = 23.36, 95 % weiblich) exemplarisch eine für die Primarstufe entwickelte Unterrichtsvideovignette mittels eigens konzipiertem Evaluationsbogen ein. Die Proband*innen schätzten dabei vorgegebene Klassenführungsstrategien und nonverbale Klassenführungskomponenten als im Video fokussiert abgebildet bzw. nicht im Video fokussiert abgebildet ein. Der Datenauswertung wurde die aus der Diagnostik bekannte Differenzierung von Sensitivität und Spezifität zugrunde gelegt (hits, misses, false alarms und correct rejections; Spinath &

Brünken, 2016). Die Prüfung von Gruppenunterschieden zwischen Lehramtsnoviz*innen und praktizierenden Lehrkräften erfolgte mittels Mann-Whitney-U-Tests.

Ergebnisse und ihre Bedeutung

Für die Klassenführungsstrategien lag die Sensitivität mit 53% (M = 0.53, SD = 0.20) leicht, jedoch signifikant, über einer Ratewahrscheinlichkeit von 50% (t(260) = 2.49, p = .014, d = 0.15). Die höhere Spezifität (M = 0.62, SD = 0.17), spricht für eine gruppenübergreifend größere Wahrscheinlichkeit, die nicht im Video abgebildeten Klassenführungsstrategien korrekt zu identifizieren. Für die nonverbalen Klassenführungskomponenten zeigte sich eine zufriedenstellende Sensitivität (M = 0.76, SD = 0.15), die Spezifität fiel jedoch deutlich geringer aus (M = 0.27, SD = 1.76). Signifikanten Gruppenunterschiede bestanden weder im Erkennen der Klassenführungsstrategien (Sensitivität: U = 2068.00, Z = -.393, p = .694; Spezifität: U = 1170.00, Z = -1.354, p = .176), noch der nonverbalen Klassenführungskomponenten (Sensitivität: U = 1668.00, Z = -1.741, p = .082; Spezifität: U = 1831.00, Z = -1.207, p = .227). Ein korrektes Erkennen scheint somit für Noviz*innen und praktizierende Lehrkräfte in hinreichendem Maße möglich – mit Einschränkungen hinsichtlich der Spezifität in Bezug auf die nonverbalen Klassenführungskomponenten. Trotz Inszenierung scheint das Erkennen dennoch hinreichend anspruchsvoll.

Im Rahmen des Symposiums werden ergänzende Befunde zur empfundenen Authentizität, Immersion und Resonanz bei der Videobetrachtung präsentiert.

Keywords: geskriptete Unterrichtsvideovignetten – (nonverbale) Klassenführung – Sensitivität & Spezifität – Authentizität, Immersion & Resonanz

Der Einfluss von Videoannotationen auf die Reflexion von geskripteten Unterrichtsvideos

Jana-Kristin von Wachter, Doris Lewalter
Technische Universität München

Theoretischer Hintergrund

Zur Schulung der professionellen Unterrichtswahrnehmung, die in mehreren Stufen von noticing, über knowledge-based reasoning und interpreting bis hin zum decision-making reicht (Jahn et al., 2014, Kaiser et al., 2017), kann die Beschäftigung mit Unterrichtsvideos einen wesentlichen Beitrag leisten. Doch oftmals werden bei einer Videoanalyse die bedeutungstragenden Sequenzen in den gezeigten Videos nicht erkannt und die anschließenden Unterrichtsreflexionen bleiben eher auf einem oberflächlichen Niveau. Daher benötigen vor allem angehende Lehrende instruktionale Unterstützung beim Umgang mit Unterrichtsvideos (Nagro et al., 2016). Auch der Übergang vom noticing zum reasoning und schließlich zum Treffen instruktionaler Entscheidungen ist ohne Hilfestellung schwierig zu bewältigen (Calandra, Gurvitch & Lund, 2008). Hier können insbesondere geskriptete Unterrichtsvideos, die typical practice abbilden, einen instruktionalen Mehrwert bieten, da sie eine komprimierte und saliente Darstellung relevanter Unterrichtsprozesse erlauben (Blomberg et al., 2013) und damit als instruktional gestaltete Videoangebote zur Verfügung stehen. Um die aktiv-produktive Auseinandersetzung mit diesen Unterrichtsvideos darüber hinaus instruktional zu unterstützen, können Videoannotationen als Reflexions- und Diskussionsanker bereitgestellt werden

(Krüger et al, 2012), der zu einer intensiveren Auseinandersetzung mit den Unterrichtsvideos anregt und deren Bearbeitung erleichtern soll.

Fragestellung

Das Ziel dieser empirischen Vergleichsstudie ist es, die Effektivität von vorgefertigten, beispielhaften Annotationen als instruktionalen Unterstützungsansatz zur Qualitätsförderung der Unterrichtsreflexion am Beispiel des thematischen Schwerpunkts motivationale Aktivierung zu ermitteln. In der Studie wird basierend auf einem Expertenrating der Videos zum Thema motivationale Aktivierung untersucht, (1) wie viele unterschiedlich saliente Unterrichtszenen in den Videos als thematisch relevant erkannt (noticing) und reflektiert (reasoning und decision making) werden und, (2) inwieweit sich diese Arbeit am Video anschließend auf die Qualität der thematisch fokussierten Gesamtreflexionen der Unterrichtsvideos abhängig vom Einsatz instruktionaler Unterstützung in der Form von vorgefertigten beispielhaften Videoannotationen auswirkt.

Methode

Die teilnehmenden Lehramtsstudierenden (N = 30, 1. - 7. Semester, gleichverteilt über Treatment- und Kontrollgruppe) betrachten drei Unterrichtsvideos aus der Toolbox Lehrerbildung, in welchen in der Testgruppe die Lernhilfen in Form von Beispielannotationen in unterschiedlicher Anzahl (fading out) vorgegeben sind. Die Kontrollgruppe erhält keine Lernhilfen. Durch die Verwendung geskripteter Videos, die bewusst thematisch relevante typical practice Beispiele in komprimierter Form enthalten, und die Bewertung durch das Expertenrating (N = 10) wird gewährleistet, dass die Videos (Dauer jeweils ca. 4-7 Min.) unterschiedlich salient ausgestaltete thematisch relevante Szenen (motivationale Aktivierung) für eine umfassende Unterrichtsreflexion beinhalten (4 hoch salient, 5 mäßig salient, 2 niedrig salient). Alle Proband*innen können in allen Videos selbstständig Annotationen erstellen. Nach dem zweiten und dritten Video verfassen sie zusätzlich eine schriftliche Reflexion mit Hilfe ihrer Annotationen.

Ergebnisse und ihre Bedeutung

Zum einen wird erwartet, dass die Lernhilfen in Form von vorgefertigten Annotationen in den Unterrichtsvideos dazu führen, dass die Teilnehmenden der Treatmentgruppe mehr und insbesondere auch niedrig saliente Szenen in den Videos kommentieren. Die bisherige Stichprobe ergab keinen signifikanten Unterschied hinsichtlich der kommentierten hoch und mäßig salienten Szenen (noticing). Jedoch zeigte sich ein tendenziell signifikanter Unterschied hinsichtlich des noticing der niedrig salienten Unterrichtsszenen zugunsten der Treatmentgruppe. Diese Befunde werden mit einer größeren Stichprobe, die durch eine weitere Erhebung im November 2022 gewonnen wird, überprüft. Zum anderen wird erwartet, dass sich die Probanden der Treatmentgruppe an den Musterannotationen orientieren und qualitativ bessere Annotationen anfertigen als die Kontrollgruppe ohne instruktionale Unterstützung. D.h., dass sie in ihren Annotationen nicht auf dem Level des noticing bleiben, sondern auch das knowledge-based reasoning/interpreting und das decision making erreichen. Die anschließende schriftliche Reflexion sollte aufbauend auf den alle Stufen der Unterrichtswahrnehmung umfassenden Annotationen ebenfalls auf einem höheren Niveau

liegen. Im Beitrag werden die Befunde präsentiert und in Hinblick auf ihre Bedeutung für die Förderung der professionellen Unterrichtswahrnehmung diskutiert.

F-14: Teacher Noticing - Konzepte, Messung, Ergebnisse und Perspektiven

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* R11 T00 D01

Teacher Noticing - Konzepte, Messung, Ergebnisse und Perspektiven

Chair(s): **Johannes König** (University of Cologne, Germany), **Gabriele Kaiser** (University of Hamburg)
*Diskutant*in(nen):* **Stefan Krauss** (Universität Regensburg)

Empirische Forschung zur professionellen Kompetenz von Lehrkräften hat in den vergangenen 15 Jahren erheblich zugenommen. Während zunächst ein besonderer Fokus auf die Modellierung und Messung von Facetten professionellen Wissens angehender (Blömeke et al., 2010a, b) wie berufstätiger Lehrkräfte (Baumert et al., 2010) gelegt wurde, erhalten seit einigen Jahren situationsspezifische Fähigkeiten, die zwischen Wissen als Disposition und unterrichtlicher Performanz von Lehrkräften vermitteln, zunehmend Aufmerksamkeit (Blömeke et al., 2015; Krauss et al., 2020). Entsprechend werden Fähigkeiten der professionellen Wahrnehmung unterrichtlicher Situationen von Lehrkräften zur Erklärung von Lernfortschritten ihrer Schülerinnen und Schüler in Modellierungen auf Unterrichtsebene inkludiert (z. B. Kersting et al., 2010; Blömeke et al., 2022). Solche Fähigkeiten der professionellen Wahrnehmung werden seit zwei Jahrzehnten im Allgemeinen als „Teacher Noticing“ empirisch untersucht, auch wenn unterschiedliche Ansätze sowie empirische Zugänge dabei erkennbar werden.

Ziel des Symposiums ist es, zunächst zum Begriff „Teacher Noticing“ bzw. zur professionellen Unterrichtswahrnehmung von (angehenden) Lehrkräften einen Überblick zum Forschungsstand zu geben sowie anschließend ausgewählte aktuelle empirische Arbeiten vorzustellen, die Einblicke über verwendete Konzepte, Messungen und Ergebnisse berichten und zur Diskussion stellen. Beitrag 1 berichtet daher über ein umfängliches, systematisches Review der Literatur zu „Teacher Noticing“, bei dem über 180 einschlägige Fachartikel der letzten zwei Jahrzehnte die Datenbasis bilden. Daraufhin folgen drei Beiträge aus der aktuellen empirischen Forschung: Beitrag 2 stellt den Einsatz videobasierter Testinstrumente aus dem TEDS-Forschungsprogramm im Rahmen der Evaluation von Schulpraktika im Rahmen des Lehramtsstudiums für Mathematik vor. Beitrag 3 stellt ein neu entwickeltes Messinstrument zur Erfassung von Entscheidungen in komplexen Klassenführungssituationen vor. Beitrag 4 präsentiert eine Untersuchung zur Entwicklung der Klassenführungskompetenzen während des ersten Berufsjahres, die als situationsspezifische Fähigkeiten mit standardisiertem Test sowie zugleich mit Selbsteinschätzungsverfahren erfasst wurden. Abschließend werden die Beiträge vom Diskutanten gewürdigt und es werden Perspektiven für zukünftige Forschung im Bereich Lehrkompetenz und „Teacher Noticing“ aufgezeigt.

Beiträge des Symposiums

Teacher Noticing – Ergebnisse einer systematischen Literaturstudie zu Konzeptualisierungen, Forschungsdesign und Studienergebnissen

Gabriele Kaiser[1], **Johannes König**[2]
[1]University of Hamburg, [2]University of Cologne

Theoretischer Hintergrund

Professionelle Unterrichtswahrnehmung von Lehrkräften (teacher noticing) ist zu einem bedeutenden Konstrukt in der Forschung zur Lehrkräfteausbildung und -fortbildung geworden, das intensiv diskutiert und beforscht wird, sowohl auf internationaler als auch nationaler Ebene (Dindyal et al.,

2021; Jahn et al., 2014). Die Ausdifferenzierung der Forschung, die im Kontext der professionellen Unterrichtswahrnehmung steht, in einzelne Studien und Ansätze verdeutlicht zugleich, dass kein einheitliches Begriffsverständnis vorliegt und der aktuelle Forschungsstand recht vielfältig ist. Vor diesem Hintergrund wurde ein umfängliches systematisches Review einschlägiger Literatur durchgeführt. Dabei standen drei Leitfragen im Mittelpunkt (König et al., 2022).

Fragestellung

1. Wie haben Forscher*innen während der letzten zwei Jahrzehnte die professionelle Wahrnehmung konzeptualisiert (einschließlich theoretischer Bezüge)?

2. Mit welchen Methoden wurde die professionelle Wahrnehmung während der letzten zwei Jahrzehnte empirisch untersucht (u.a. Studiendesigns)?

3. Welche Erkenntnisse wurden dabei erzielt, u.a. bezogen auf den Erwerb professioneller Wahrnehmung im Rahmen der Lehrer*innenbildung?

Methode

Es erfolgte eine systematische Literatursuche in internationalen hochkarätigen Zeitschriften. Grundlage für die Datenanalyse stellt eine Volltextanalyse von 182 Artikeln dar, die in renommierten Datenbanken und begutachteten englischsprachigen Fachzeitschriften veröffentlicht wurden. Die Texte wurden anhand eines entlang der drei Fragen entwickelten, umfänglichen Systems mit 58 Kategorien von geschulten Ratern kodiert. Die Interrater-Reliabilität von 20% der doppelt kodierten Artikel war gut (durchschnittliches Kappa 0,72).

Ergebnisse und Bedeutung

Hinsichtlich der Frage 1 zeigt sich, dass der Diskurs international dominiert wird vom Konstrukt „noticing“, wobei insbesondere im deutschsprachigen Raum der naturwissenschaftlichen Didaktiken das Konstrukt von „professional vision“ zentral ist. Theoretische Basis der aktuellen Diskussion zur professionellen Unterrichtswahrnehmung sind verschiedene Richtungen der professionellen Unterrichtswahrnehmung, die den aktuellen Diskurs ausgeschärft haben:

• eine kognitionspsychologische Perspektive, die das Wahrnehmen von Lehrkräften als das definiert, worauf Lehrkräfte achten und wie sie es interpretieren, wobei der Schwerpunkt auf den entsprechenden mentalen Prozessen liegt (van Es & Sherin, 2002);

• eine soziokulturelle Perspektive des professionellen Wahrnehmens von Lehrkräften, die sich auf die soziale und situierte Natur des Wahrnehmens von Lehrkräften konzentriert und oft in Verbindung mit dem Begriff "professional vision" erwähnt wird (Goodwin, 1994);

• eine disziplinspezifische Perspektive auf die professionelle Unterrichtswahrnehmung von Lehrkräften, die sich auf die professionelle Förderung von Praktiken zur Sensibilisierung von Lehrkräften konzentriert (Mason, 2002);

• eine auf das Konstrukt Expertise bezogene Perspektive der professionellen Unterrichtswahrnehmung, die die Unterschiede zwischen Expert*innen und Noviz*innen fokussiert (Berliner, 2001).

Die Ergebnisse der durchgeführten Literaturanalyse zeigen die Dominanz einer kognitionspsychologischen Perspektive auf die professionelle Unterrichtswahrnehmung von Lehrkräften auf. Obwohl die professionelle Unterrichtswahrnehmung von Lehrkräften in früheren Arbeiten als ganzheitliches Konzept beschrieben wurde, tendieren neuere Studien aus dem letzten Jahrzehnt dazu, die professionelle Unterrichtswahrnehmung in verschiedene Facetten zu differenzieren, nämlich Wahrnehmung, Interpretieren und ggf. auch Entscheiden über den weiteren Unterrichtsgang (Blömeke et al., 2015; Kaiser et al., 2015).

Mit Bezug auf Frage 2 zeigen die Ergebnisse u.a., dass insgesamt qualitative Studien mit kleinen Stichprobengrößen vorherrschend sind, was zwar eine feinkörnige Analyse ermöglicht, aber die Verallgemeinerbarkeit der Ergebnisse einschränkt. Zugleich werden die Grenzen der aktuellen Diskussion erkennbar, wie z.B. eine Dominanz von Studien zur professionellen Unterrichtswahrnehmung von Lehrkräften vor allem im Bereich der Mathematik. Des Weiteren ist die geringe Repräsentanz von Arbeiten aus anderen Teilen der Welt als Nordamerika und Europa auffällig.

Schließlich werden mit Blick auf Frage 3 die in dieser Literaturübersicht analysierten Studien im Hinblick auf ihre Ergebnisse zur Verbesserung der professionellen Unterrichtswahrnehmung von Lehrkräften bezogen auf Vergleiche zwischen Expert*innen und Anfänger*innen eingehend untersucht, womit Perspektiven zum Stellenwert von professioneller Unterrichtswahrnehmung sowohl in der Ausbildung von Lehrkräften als auch in der Fortbildung von Lehrkräften aufgezeigt werden.

Einfluss (verlängerter) schulischer Praxisphasen auf die professionelle Unterrichtswahrnehmung

Anton Bastian[1], Jonas Weyers[2], Johannes König[2], Gabriele Kaiser[1]
[1]University of Hamburg, [2]University of Cologne

Aufgrund des bedeutenden, empirisch belegten Einflusses von Lehrkräften auf den Lernerfolg von Schüler*innen ist die Lehramtsausbildung im Fokus der Bildungspolitik und empirischen Bildungsforschung (z. B. Blömeke et al., 2022; Hattie, 2009). Als Antwort auf Disparitäten zwischen akademischer Lehrkräfteausbildung und Berufspraxis, auch als Theorie-Praxis-Problem bezeichnet (König & Rothland, 2018; Korthagen, 2010), wurden in vielen Bundesländern Deutschlands in den vergangenen Jahren verlängerte Phasen schulischer Praxis in die erste Phase der Lehramtsausbildung integriert, um früher berufspraktische Lerngelegenheiten zu bieten sowie Theorie und Praxis des Lehrberufs zu vernetzen (Ulrich & Gröschner, 2020). Jedoch liegen zu den Wirkungen dieser Praxisphasen bislang nur wenige empirische Untersuchungen vor, die über Selbsteinschätzungen von Studierenden hinausgehen (Ulrich et al., 2020).

Ein wichtiges Ziel der Praxisphase ist die Kontextualisierung, die situationsspezifische Anwendung und somit Prozeduralisierung von im Studium vermittelten professionellen Wissen (König, 2019). Diese Vernetzung von dispositional angelegtem Wissen und Performanz im Unterricht wird u.a. durch die Fähigkeiten der professionellen Unterrichtswahrnehmung gefördert. Konzeptualisiert durch die Facetten Wahrnehmen, Interpretieren und Entscheidungentreffen beschreibt dieses Konstrukt die situationsspezifischen Fähigkeiten von Lehrkräften, die zwischen Wissen und Performanz mediieren und damit essenzielle Bestandteile der professionellen Kompetenz von Lehrkräften darstellen

(Blömeke et al., 2015; Dindyal et al., 2021; Kaiser et al., 2015). Allerdings ist bisher nur wenig untersucht, inwieweit Praxisphasen Fähigkeiten der professionellen Unterrichtswahrnehmung bei Lehramtsstudierenden ausbilden, insbesondere in quantitativen Längsschnittstudien (Ulrich et al., 2020). Eine Ausnahme bildet die Studie von Mertens und Gräsel (2018), welche positive Effekte einer fünfmonatigen Praxisphase auf allgemeinpädagogische Facetten der professionellen Unterrichtswahrnehmung nachweisen.

Im Rahmen des Projekts TEDS-Validierung-Transfer, eines Folgeprojekts der Teacher Education and Development Study in Mathematics (TEDS-M), beschäftigt sich unsere Studie mit der Frage nach Effekten von schulischen Praxisphasen auf die professionelle Unterrichtswahrnehmung von Mathematiklehramtsstudierenden sowohl unter einer allgemeinpädagogischen als auch einer mathematikdidaktischen Perspektive. Des Weiteren betrachten wir in den vorgestellten Analysen differenzielle Effekte verschiedener potenzieller Einflussfaktoren, also Input- und Prozessfaktoren (Ulrich et al., 2020), die Rahmenbedingungen der Praxisformate auf die professionelle Unterrichtswahrnehmung, wie Lerngelegenheiten, lernprozessbezogene Tätigkeiten und mentorielle Unterstützung darstellen.

Zu diesem Zweck wurden die professionelle Unterrichtswahrnehmung von n = 174 Lehramtsstudierenden unterschiedlicher Lehrämter von sechs Universitäten verschiedener Bundesländer und Rahmenbedingungen ihrer Praxisphasen im Masterstudium längsschnittlich im Pre-Post-Design untersucht. Für die Messung der professionellen Unterrichtswahrnehmung wurden zwei bewährte, videogestützte Instrumente, bezogen auf das Grundschul- bzw. Sekundarstufenlehramt eingesetzt (siehe Kaiser et al., 2015), Rahmenbedingungen wie Lerngelegenheiten wurden mit geschlossenen Items, insbesondere Ratingskalen, erhoben (z. B. Doll et al., 2018). Die Daten der Tests zur professionellen Unterrichtswahrnehmung wurden auf Basis der Item Response Theory mit Raschmodellen skaliert. Anschließend wurde die Frage nach dem Einfluss der Praxisphasen auf die professionelle Unterrichtswahrnehmung regressionsanalytisch im Rahmen von Strukturgleichungsmodellen bearbeitet.

Die Ergebnisse zeigen einen signifikanten Anstieg ($p < .001$) der Fähigkeiten in allen drei Facetten der professionellen Unterrichtswahrnehmung mit kleinen Effektstärken. Die Autorkorrelation liegt mit Werten zwischen .50 und .60 im großen Effektbereich. Erste Regressionen mit großer, signifikanter Varianzaufklärung ($R^2 > .25$) belegen die Einflüsse der Abiturnote und des Beziehens pädagogischer und mathematikdidaktischer Theorien auf konkrete Situationen auf die Interpretationsfähigkeiten. Diese Einflüsse finden sich nicht für die anderen Facetten der professionellen Unterrichtswahrnehmung und weisen auf die besondere Bedeutung der Interpretation unterrichtlichen Situationen und ihrer starken Verknüpfung zu kognitiven Dispositionen hin. Des Weiteren deutet sich ein Einfluss der Interpretationsfähigkeiten auf das Erlernen des Entscheidungentreffens an.

Insgesamt deuten die Ergebnisse auf den gewichtigen Einfluss von berufspraktischen Lerngelegenheiten in der ersten Phase der Lehramtsausbildung auf die Entwicklung der professionellen Unterrichtswahrnehmung hin und betonen damit erneut die Relevanz von Praxiserfahrungen für die Ausbildung situationsspezifischer Fähigkeiten (Bastian et al., 2022). Der

Zusammenhang zwischen Interpretationsfähigkeiten und dem Beziehen von Theorien auf Situationen stützt den Nutzen entsprechender Begleitveranstaltungen für Praxisphasen und legt die Ausweitung solcher Angebote in diesem Bereich nahe.

Decision-Making im Classroom Management standardisiert messen - Zum Einsatz eines videobasierten Testinstruments in der Lehrkräfteausbildung

Jonas Weyers, Charlotte Kramer, Kai Kaspar, Johannes König
University of Cologne

Theoretischer Hintergrund

Jüngere Konzeptionen von Lehrkräftekompetenz berücksichtigen neben dem professionellen Wissen auch die situationsspezifischen Fähigkeiten, d. h. wissensgeleitete Wahrnehmung, Interpretation und Entscheidungsfindung im Unterricht (Blömeke et al., 2015; Kaiser et al., 2015; König et al., 2022). Mit der Ausweitung schulpraktischer Lerngelegenheiten innerhalb der universitären Lehrkräfteausbildung (z. B. Christiansen & Erixon, 2021; Ulrich & Gröschner, 2020) wächst auch das Anliegen, situationsspezifische Fähigkeiten als Lernergebnis der Lehrkräfteausbildung zu erfassen. Dabei werden vermehrt videobasierte Testinstrumente eingesetzt (Kramer et al., 2017; Seidel & Stürmer, 2014; Wiens et al., 2013).

Besonders deutlich wird die Notwendigkeit der Berücksichtigung situationsspezifischer Kompetenzen im Bereich Classroom Management, verstanden als lernförderliche Gestaltung von Lernumgebungen (Evertson & Weinstein, 2006). Zur Durchführung effektiven Classroom Managements müssen Lehrkräfte kritische Momente – beispielsweise Unterrichtsstörungen – erkennen und interpretieren sowie innerhalb der Unterrichtssituation angemessene Entscheidungen treffen.

Während zur Wahrnehmung und Interpretation von Classroom Management bereits erprobte Messinstrumente vorliegen (Gold & Holodynski, 2017; Jamil et al., 2015; König, 2015), wird die Entscheidungsfindung, nachfolgend „Decision-Making“, im Kontext standardisierter Kompetenzmessung selten adressiert. Dieser Beitrag fokussiert das Decision-Making als Teil der Classroom Management Expertise (CME) und präsentiert ein Instrument zur Erfassung dieser Kompetenzfacette, den CME-Decide-Test. Dieser ergänzt den bestehenden CME-Test mit Schwerpunkt auf Wahrnehmung und Interpretation (König, 2015; König & Lebens, 2012).

Fragestellung

Der Beitrag nimmt den Einsatz des CME-Decide-Tests in der Lehrkräfteausbildung unter drei Fragestellungen in den Blick. Fragestellung 1 fokussiert die Reliabilität der Messung mit Lehramtsstudierenden. Fragestellung 2 prüft die Korrelationen zu assoziierten Konstrukten, konkret zum pädagogischen Wissen sowie zur Wahrnehmung und Interpretation im Classroom Management. Die Validität der Testwertinterpretation als Lernergebnis der Lehrkräfteausbildung steht im Zentrum von Fragestellung 3, wobei Korrelationen zu bildungswissenschaftlichen Ausbildungsinhalten und schulpraktischen Lerngelegenheiten untersucht werden.

Methode

Der CME-Decide-Test besteht aus 12 kurzen Videoclips (Videodauer < 1 min) mit Unterrichtssituationen, in denen zentrale Anforderungen des Classroom Managements abgebildet

sind, z. B. Unterrichtsstörungen und Klassenregeln. Nach jedem Videoclip wird in einem offenem Antwortformat abgefragt, wie sich die Testperson in der Rolle der Lehrkraft im Anschluss an die gesehene Unterrichtssituation verhalten würde. Das Scoring erfolgte auf Basis eines Kodiermanuals. Das Instrument wurde im Rahmen des Bildungsmonitorings im Projekt Zukunftsstrategien Lehrer*innenbildung (ZuS) der Universität zu Köln in den Jahren 2019 (n = 136) und 2020 (n = 150) mit Lehramtsstudierenden eingesetzt.

Die Prüfung der Reliabilität und Dimensionalität erfolgte unter Einbezug von Strukturgleichungs- und Raschmodellierungen. Zur Untersuchung der Validität von Testwertinterpretationen wurden die Korrelationen zu etablierten Kompetenztests zum pädagogischen Wissen (König & Blömeke, 2010) sowie zur Wahrnehmung und Interpretation im Bereich Classroom Management (König, 2015) herangezogen. Ferner wurden Korrelationen zu schulpraktischen und universitären Lerngelegenheiten geprüft.

Ergebnisse und ihre Bedeutung

In vorläufigen Analysen zeigt der CME-Decide-Test geringe – für weitere Analysen jedoch akzeptable – Reliabilität ($\alpha \geq .50$), was die Herausforderungen der Messung von Decision Making unterstreicht (vgl. Kaiser et al., 2015; Shavelson, 2010). Es zeigt sich ein geringer Zusammenhang zum pädagogischen Wissen ($r = .15$) sowie ein moderater Zusammenhang zur Wahrnehmung/Interpretation im Classroom Management ($r = .40$). Das Decision-Making ist somit von assoziierten Konstrukten deutlich abgrenzbar. Insbesondere die Zusammenhänge zur Wahrnehmung bzw. Interpretation stehen mit theoretischen Vorannahmen in Einklang, während der geringe Zusammenhang zum pädagogischen Wissen angesichts der Annahme wissensbasierter Entscheidungsfindung kritisch zu diskutieren ist.

Ferner sind geringe Korrelationen ($r = .20$) zu praxisbezogenen Lerngelegenheiten (z. B. das Beobachten von Unterricht bzw. das Beziehen von Theorien auf Unterrichtssituationen) festzustellen. Demgegenüber zeigen universitäre Lerngelegenheiten, welche als bildungswissenschaftliche Inhaltsbereiche operationalisiert wurden, keine Effekte. Der Befund legt nahe, dass der Erwerb situationsspezifischer Fähigkeiten im Rahmen schulpraktischer Phasen des Lehramtsstudiums förderbar ist (vgl. Mertens & Gräsel, 2018; Stürmer et al., 2013). Die Ergebnisse werden zusammenfassend mit Blick auf den Einsatz des CME-Decide-Tests für Forschung und Evaluation sowie in Bezug auf Herausforderungen der videobasierten Kompetenzmessung diskutiert.

Zur Entwicklung von selbst wahrgenommenen und standardisiert gemessenen Klassenführungskompetenzen beim Einstieg in den Lehrberuf

Christian Brühwiler[1], Lena Hollenstein[1], Manuela Haldimann[1], Doreen Flick-Holtsch[2]
[1]Pädagogische Hochschule St.Gallen, [2]Universität Zürich

Theoretischer Hintergrund

Während des Berufseinstiegs erleben Lehrpersonen den Unterricht als besonders herausfordernd (Doyle, 2006; Wanzare, 2007). Studien haben wiederholt gezeigt, dass beim Übergang in die Berufstätigkeit Themen der Klassenführung (z.B. der Umgang mit störendem Verhalten) eine der größten Herausforderungen für Lehrpersonen darstellen (Terhart, 1994; Zingg, & Grob, 2002).

Eine der Herausforderungen besteht darin, dass sich Studierende während des Studiums zwar viel Wissen über Klassenführung aneignen, sie aber nicht notwendigerweise Klassenführungskompetenzen erwerben (Jones, 2006; Stough et al., 2015). Bezogen auf die Klassenführung erscheinen situationsspezifische Fähigkeiten besonders bedeutsam zu sein (Jones, 2006). Situationsspezifische Fähigkeiten gelten als situativ, intuitiv und implizit (Leinhardt et al., 1995) und können als verhaltensnahe Indikatoren von Lehrpersonenexpertise gesehen werden. Sie umfassen die Wahrnehmung und die Interpretation von lernrelevanten Situationen sowie das anschließende Entscheiden über Handlungsoptionen (Blömeke et al., 2015; Kramer et al., 2017). Forschungsvorhaben, die vor allem Wahrnehmung und Interpretation fokussieren, sind häufig unter Begriffen wie professionelle Unterrichtswahrnehmung (engl.: professional vision) oder Noticing zu finden (Seidel & Stürmer, 2014; Sherin et al., 2011).

Es ist davon auszugehen, dass sich Klassenführungskompetenzen während des ersten Berufsjahres aufgrund unterschiedlicher formeller und informeller Lerngelegenheiten aufbauen und weiterentwickeln (König et al., 2017; Schmidt et al., 2011). So wiesen König und Kramer (2016) nach, dass sich die Klassenführungskompetenzen während des Referendariats verbessern. Die Forschung zum Übergang von der Ausbildung in den Lehrberuf ist insgesamt eher spärlich. Kaum eine Studie untersucht längsschnittlich, wie sich Klassenführungskompetenzen während des Berufseintritts verändern.

Fragestellung

In diesem Beitrag soll die Entwicklung der Klassenführungskompetenzen, sowohl gemessen mittels standardisiertem Test als auch mittels Selbsteinschätzung, während des ersten Berufsjahres untersucht werden. Folgende Fragestellungen stehen im Zentrum:

1. Über welche Klassenführungskompetenzen verfügen angehende Lehrpersonen am Ende des Studiums?

2. Wie entwickeln sich selbsteingeschätzte und standardisiert gemessene Klassenführungskompetenzen von Lehrpersonen während des ersten Berufsjahrs?

3. Wie hängen Lerngelegenheiten im Bereich Klassenführung mit der Entwicklung von Klassenführungskompetenzen im ersten Berufsjahr zusammen?

Methode

Als Datengrundlage dient das Evaluationsprojekt «Evaluierung und Weiterentwicklung der Pädagog*innenbildung NEU in Österreich». Insgesamt liegen längsschnittliche Daten von 96 (angehenden) Lehrpersonen der Primarstufe und Sekundarstufe Allgemeinbildung am Ende des Studiums sowie im ersten Berufsjahr vor. Die Klassenführungskompetenzen wurde einerseits mit einem Testinstrument zur Erfassung pädagogisch-psychologischen Handlungswissens zur Klassenführung (PPHW-K; Brühwiler et al., 2017) gemessen. Der PPHW-K ist ein textvignettenbasierter Test, der die Erfassung der situationsspezifischen Fähigkeiten Interpretieren und Entscheiden fokussiert (max. zu erreichende Punktzahl = 28; MZP1: M = 18.5; SD = 4.26; MZP2: M = 18.0; SD = 4.13). Die Kompetenzselbsteinschätzungen der (angehenden) Lehrpersonen wurden

auf einer 6-stufigen Likert-Skala erfasst (Beispielitem: «Ich kann verhindern, dass einige Schüler*innen meine Unterrichtsstunde massiv stören.»; MZP1: M = 4.53; SD = .97; MZP2: M = 4.86; SD = 0.87). Die Lerngelegenheiten wurden dichotom erfasst. Die (angehenden) Lehrpersonen wurden zu den Items der Kompetenzselbsteinschätzung jeweils gefragt, wo sie diese Inhalte gelernt haben (Hochschullehrveranstaltung oder Praktika bzw. Mentoring in der Induktionsphase oder im Unterricht; Mehrfachantworten waren möglich).

Ergebnisse und Bedeutung

Erste Ergebnisse weisen auf einen Anstieg der Kompetenzselbsteinschätzung im ersten Berufsjahr hin ($t(91) = -3.54$, $p = .001$, $d = .36$). Kein signifikanter Unterschied zeigt sich hingegen bei der Erfassung mit dem Testinstrument PPHW-K ($t(69) = 0.08$, $p = .94$, $d = .08$). Ein positiver Zusammenhang im ersten Berufsjahr besteht zwischen den Lerngelegenheiten während dem Mentoring und der Induktionsphase ($r = .21$, $p = .06$) sowie der Kompetenzselbsteinschätzung und dem Unterricht ($r = .20$, $p = .05$).

Die Ergebnisse werden in Bezug auf die Bedeutung der Lerngelegenheiten für die Entwicklung von Klassenführungskompetenzen (Selbsteinschätzung sowie situationsspezifische Fähigkeiten im Bereich Klassenführung) im ersten Berufsjahr diskutiert.

F-15: Kompetenzentwicklung in der beruflichen Ausbildung: Prädiktoren und Outcomes differenzieller Lernverläufe

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* T03 R03 D75

Kompetenzentwicklung in der beruflichen Ausbildung: Prädiktoren und Outcomes differenzieller Lernverläufe

Chair(s): **Julian M. Etzel** (IPN - Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany), **Jennifer Schauer** (Technische Universität Dresden)
*Diskutant*in(nen):* **Andreas Frey** (Goethe-Universität Frankfurt)

Das zentrale Ziel der Berufsausbildung ist der Erwerb berufsspezifischer Fähigkeiten und Fertigkeiten, die Auszubildende dazu befähigen sollen, ihre spätere berufliche Tätigkeit qualifiziert und selbständig auszuüben (Minnameier, 2013). Darüber hinaus existieren sekundäre Ziele, die sich als Resultate von Sozialisationsprozessen manifestieren. Hierzu zählen unter anderem die Identifikation mit der beruflichen Umwelt oder die Entwicklung übergreifender Kompetenzen als Voraussetzungen zur gesellschaftlichen Teilhabe (z.B. Lempert, 2006).

Neben der beruflichen *Kompetenzmodellierung* und *-diagnostik* widmet sich die berufliche Bildungsforschung zunehmend der empirischen Untersuchung der *Kompetenzentwicklung* (z.B. Winther, 2010; Seeber & Nickolaus, 2010). Hauptaugenmerk liegt dabei auf der Entwicklung beruflicher Fachkompetenzen (Rüschoff, 2019). Im Rahmen von *Kompetenzerklärungsmodellen* stehen zudem Entwicklungsbedingungen auf individueller Ebene – d.h. soziale, kognitive und non-kognitive Merkmale von Auszubildenden – sowie auf organisationaler Ebene – d.h. Merkmale der betrieblichen und schulischen Lernorte und ihrer Wahrnehmung durch die Auszubildenden – im Fokus (z.B. Abele, 2014; Baethge-Kinsky, Baethge & Lischewski , 2016).

Das vorliegende Symposium beleuchtet offene Forschungsfragen zur Entwicklung von Auszubildenden in der Berufsausbildung aus psychologischer, berufs- und wirtschaftspädagogischer sowie bildungssoziologischer Perspektive. Die vier längsschnittlichen empirischen Untersuchungen fokussieren relevante Entwicklungsresultate und -bedingungen in unterschiedlichen Phasen des Ausbildungsverlaufs.

Im ersten Beitrag wird die Rolle der allgemeinen kognitiven Fähigkeiten auf die Entwicklung berufsfachlicher Kompetenzen in Ausbildungsgängen mit technisch-naturwissenschaftlichem und mathematischem Schwerpunkt untersucht. Zentrale Fragestellungen des Beitrags sind ob sich differenzielle Leistungsentwicklungen in Abhängigkeit der jeweiligen Lernumwelt ergeben und inwieweit diese differenziellen Entwicklungsverläufe von interindividuellen Unterschieden in kognitiven Fähigkeiten abhängen.

Der zweite Beitrag untersucht die Entwicklung berufsfachlicher Kompetenzen in der zweiten Ausbildungshälfte am Beispiel von Kfz-Mechatroniker:innen. Im Zentrum steht die Frage, welche Bedeutung dem Niveau der schulischen Vorbildung zukommt, mit dem sich die Auszubildenden in die Berufsausbildung eingliedern, und inwiefern sich die Kompetenzentwicklung in Abhängigkeit des allgemeinbildenden Schulabschlusses unterscheidet.

Im dritten Beitrag steht die Kompetenz zur Bewältigung von Belastungen und Anforderungen von Auszubildenden in Pflegeberufen im Fokus. Ziel der Arbeit ist die Untersuchung der Entwicklung der Belastung-Bewältigungs-Kompetenz von angehenden Pflegekräften sowie die Beantwortung der

Frage, welchen Einfluss Arbeits- und Interaktionsmerkmale als potenzielle Stützfaktoren auf die Entwicklung haben.

Der vierte Beitrag widmet sich abschließend der organisationalen Identifikation von Auszubildenden in den ersten Monaten der Berufsausbildung. Neben der Entwicklung der organisationalen Identifikation werden mögliche Prädiktoren dieses Prozesses im Zusammenspiel mit Aufgabenbewältigung und Engagement untersucht. Weiterhin wird untersucht, welche Konsequenzen die organisationale Identifikation für die Leistung der Auszubildenden hat.

Beiträge des Symposiums

Die Rolle kognitiver Fähigkeiten für differenzielle Entwicklungsverläufe berufsfachlicher Kompetenzen

Julian M. Etzel, Gabriel Nagy
IPN - Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany

Das zentrale Ziel der Berufsausbildung ist die Vermittlung berufsfachlicher Kompetenzen (Minnameier, 2013). Wie erfolgreich der Kompetenzerwerb in der Ausbildung jedoch tatsächlich ist, wurde bis zum heutigen Tage nur unzureichend empirisch untersucht (Seeber & Lehmann, 2011). So gibt es zum Beispiel kaum empirische Forschungsarbeiten zu der Frage inwieweit sich domänenspezifische Leistungszuwächse zwischen verschiedenen Ausbildungsgängen unterscheiden. Dies ist überraschend, da gerade solche Untersuchungen belastbare Hinweise für die Effektivität des Kompetenzerwerbs in der Berufsausbildung liefern können.

Gegenstand der vorliegenden Arbeit ist die Untersuchung der Leistungsentwicklung in Ausbildungsberufen mit technisch-naturwissenschaftlichen und mathematischen Schwerpunkten sowie die Untersuchung der Rolle der kognitiven Grundfähigkeit als potenziellem Prädiktor der Leistungsentwicklung. Betrachtet werden Leistungen in den Domänen Physik, Chemie und Mathematik, die mit berufsfachlichen Leistungstests erfasst wurden. Die verwendeten Leistungstests erfassen dabei explizit die Fähigkeit zur Anwendung schulisch erworbenen Wissens auf berufsspezifische Problemstellungen. Da sich die Aufgaben der Leistungstests an den Inhalten der Ausbildungsberufe orientieren (Mathematik: Industriekaufleute, Chemie: Chemielaborant:innen und Physik: gewerblich-technische Berufe) stellen sie proximale Indikatoren des berufsfachlichen Kompetenzerwerbs dar.

Die vorliegende Arbeit zielt auf zwei zentrale Forschungsfragen zur Kompetenzentwicklung in der Berufsausbildung ab. Erstens wird untersucht, ob sich ein Effekt des Ausbildungskontextes auf die Entwicklung berufsfachlicher Leistungen nachweisen lässt. Zweitens wird sich der Frage gewidmet, welche Rolle die kognitiven Fähigkeiten der Auszubildenden für die berufsfachliche Leistungsentwicklung haben. Bei der Untersuchung der ersten Fragestellung dienen die unterschiedlichen Ausbildungsgruppen als Treatment- bzw. Kontrollgruppen. Ein positiver Effekt des Ausbildungskontextes liegt vor, wenn die domänenspezifische Leistungszuwächse mit der inhaltlichen Ausrichtung der jeweiligen Ausbildungsgänge korrespondieren. So sollten, zum Beispiel, Chemielaboranten in der Fachdomäne Chemie die größten Leistungszuwächse erzielen, wohingegen für Auszubildende anderer Ausbildungsgruppen keine solchen Zuwächse zu erwarten sind. Im Hinblick auf die zweite Fragestellung wird die kognitive Grundfähigkeit, die sich als starker Prädiktor

von allgemeinem Bildungserfolg erwiesen hat (z.B. Deary et al., 2007), als ein plausibler Prädiktor für Leistungszugewinne untersucht. Dabei stellt sich insbesondere die Frage, ob potenzielle Effekte der kognitiven Grundfähigkeit lediglich auf das Ausgangsniveau der berufsfachlichen Leistungen beschränkt sind, oder ob Auszubildende mit höheren kognitiven Fähigkeiten die Lerngelegenheiten der dualen Ausbildung besser nutzen können und somit größere Leistungsgewinne erzielen.

Datengrundlage der Arbeit bildete die ManKobE-Studie (Retelsdorf et al., 2013), in deren Rahmen Auszubildende aus verschiedenen naturwissenschaftlich-mathematischen Ausbildungsgängen über den gesamten Verlauf der Ausbildung an drei Messzeitpunkten untersucht wurden. Die Analysestichprobe bildeten N = 2,771 Auszubildende (nGWT = 1,546, nIK = 528, nCL= 697). Die berufsfachlichen Leistungen in den drei Domänen Physik, Mathematik und Chemie wurden mit Hilfe von eigens für die Studie entwickelten Aufgaben erfasst. Diese können mit schulisch erworbenem Wissen gelöst werden, sind jedoch explizit in den jeweiligen beruflichen Kontext eingebettet. Die kognitiven Grundfähigkeiten wurden mittels numerischer, figuraler und verbaler Subskalen des Kognitiven Fähigkeitstests (KFT; Heller & Perleth, 2000) erfasst.

Die Ergebnisse der Analysen zeigen, dass die höchsten Leistungszuwächse im beruflichen Fachwissen in denjenigen Fachdomänen zustande kommen, die im Fokus des jeweiligen Ausbildungsganges stehen. Chemielaboranten zeigten die höchsten Leistungszuwächse in Chemie, Industriekaufleute in Mathematik und Auszubildende in technischen Ausbildungsberufen in Physik. In den fachfremden Domänen (z.B. Chemie bei den technischen Ausbildungsberufen) zeigte sich im Durchschnitt keine Entwicklung. Einzige Ausnahme bildete dabei die Physikleistung (signifikanter Zuwachs bei Chemielaboranten, signifikanter Rückgang bei Industriekaufleuten). Während Auszubildende mit höheren kognitiven Grundfähigkeiten ein höheres Ausgangsleistungsniveau aufwiesen, ergaben die Analysen keine Hinweise auf einen Effekt der kognitiven Grundfähigkeiten auf den Leistungszuwachs. Zusammenfassend belegen die Befunde die Effektivität der Berufsausbildung für die Entwicklung berufsfachlicher Kompetenz und zeigen, dass höhere kognitive Fähigkeiten mit dem Niveau der Leistungen zu Beginn der Ausbildung assoziiert sind, nicht jedoch mit interindividuellen Unterschieden in der Leistungsentwicklung.

Entwicklung berufsfachlicher Kompetenzen von Auszubildenden – Welche Rolle spielt der Schulabschluss?

Jennifer Schauer[1], Stephan Abele[1], Julian M. Etzel[2]
[1]Technische Universität Dresden, [2]IPN - Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany

In den Ausbildungschancen von Jugendlichen bestehen fortwährend Ungleichheiten nach schulischem Vorbildungsniveau: Jugendliche mit max. Hauptschulabschluss (HSA) münden deutlich seltener in eine Berufsausbildung ein als jene mit Mittlerem Schulabschluss (MSA) oder (Fach-)Hochschulzugangsberechtigung (HZB; zuletzt AGBB, 2022). Neben an die schulische Vorbildung geknüpfte Zugangsvoraussetzungen des Schulberufssystems werden Schulabschlüsse in betrieblichen Rekrutierungsprozessen als Leistungs- und Entwicklungsindikatoren bzw. Produktivitätssignale herangezogen (Granato & Ulrich, 2014). Insbesondere bei Jugendlichen mit max. Hauptschulabschluss (HSA) wird hierbei ‚mangelnde Ausbildungsreife' wahr- und angenommen. Damit sind kognitive und non-kognitive Defizite gemeint (z.B. in schulischen Basiskompetenzen,

Merkmalen des Arbeits- und Sozialverhaltens), die den Ausbildungserfolg gefährden können (Eberhard & Ulrich, 2013). Studien zeigen zwar Bildungsgruppenunterschiede in den mittleren Ausprägungen dieser Merkmale. Darüber hinaus weisen die Bildungsgruppen allerdings eine große Binnenheterogenität auf und die Kompetenz- und Fähigkeitsausprägungen überlagern sich teilweise (z.B. Holtmann et al., 2018 für Jugendliche mit max. HSA und MSA).

Eine zentrale Dimension von Ausbildungserfolg ist der Erwerb berufsfachlicher Kompetenzen. Ob und inwiefern sich Auszubildende, die mit unterschiedlichem Schulabschluss in die Berufsausbildung einmünden, in ihrer berufsfachlichen Kompetenzentwicklung unterscheiden, wurde bislang wenig erforscht. In unserem Beitrag untersuchen wir diese Bildungsgruppenunterschiede am Beispiel von Kfz-Mechatroniker:innen. Wir nutzen Daten der Längsschnittstudie ManKobE (Mathematisch-Naturwissenschaftliche Kompetenzen in der beruflichen Erstausbildung; Retelsdorf et al., 2013). Die Untersuchungsstichprobe der Kfz-Mechatroniker:innen umfasst eingangs NT1 = 823 Auszubildende. Berufsfachliche Kompetenzen wurden entlang der Dimensionen Fachwissen und Problemlösekompetenz im zweiten und dritten Ausbildungsjahr (2. und 3. AJ) erfasst. In Strukturgleichungsmodellen (v.a. Latent Change Analysen) untersuchen wir die Entwicklung des Fachwissens in der zweiten Ausbildungshälfte unter Berücksichtigung des Schulabschlussniveaus der Auszubildenden (NT2,T3 = 433).

Unsere Analysen zeigen einen signifikanten Fachwissenszuwachs in der Gesamtgruppe (ΔMW = .275, p = .012). Dieser korreliert positiv mit dem individuellen Vorwissensniveau im 2. AJ (βT2 = .262, p < .001). Der Fachwissenszuwachs in der zweiten Ausbildungshälfte ist also insbesondere für die bereits Leistungsstärkeren groß. Mit Blick auf die Bildungsgruppen finden wir signifikante Unterschiede im Vorwissensniveau: Im 2. AJ zeigen Auszubildende mit MSA (βMSA,T2 = .699, p < .001) und (Fach-)HZB (βHZB,T2 = 1.102, p < .001) ein höheres Fachwissensniveau als jene mit max. HSA. Unter Kontrolle dieser Vorwissensunterschiede finden wir keinen weiteren Bildungsgruppeneffekt auf den Fachwissenszuwachs (βMSA,ΔMW = -.135, p = .332; βHZB,ΔMW = -.146, p= .548).

Für die Fachwissensentwicklung von Kfz-Auszubildenden in der zweiten Ausbildungshälfte scheint also das individuelle Vorwissen der Auszubildenden entscheidend – je höher das Ausgangsniveau, desto höher der Fachwissenszuwachs. Mit Blick auf die Bildungsgruppenunterschiede deuten unsere Analysen persistente bzw. zunehmende mittlere Fachwissensunterschiede an: Da Auszubildende mit MSA und (Fach-)HZB im 2. AJ über ein durchschnittlich höheres Vorwissen verfügen als jene mit max. HSA, sind für sie höhere Zuwächse zu erwarten. Gleichzeitig deuten die Analysen darauf hin, dass sich Auszubildende mit gleichem Vorwissensniveau unabhängig vom Schulabschluss entwickeln. Schulabschlüsse sind daher nur bedingt valide Indikatoren der individuellen Entwicklung in der Berufsausbildung. Die Frage der Generalisierbarkeit unserer Befunde für die Problemlösekompetenz von Kfz-Auszubildenden sowie weitere Ausbildungsberufe der ManKobE-Stichprobe ist Ausgangspunkt weiterführender Analysen.

Belastungsbewältigung als Ausbildungsziel in der Pflege – Verlauf und Stützfaktoren des Kompetenzerwerbs

Julia Warwas[1], Wiebke Vorpahl[1], Susan Seeber[2], Philine Krebs[2], Ulrike Weyland[3], Eveline Wittmann[4], Larissa Wilzeck[3], Aldin Strikovic[4], Svenja Hill[4]
[1]Universität Hohenheim, [2]Georg-August-Universität Göttingen, [3]Westfälische Wilhelms-Universität Münster, [4]Technische Universität München

Das komplexe Konstrukt beruflicher Handlungskompetenz, auf das die berufliche Ausbildung im jeweiligen Kontext abzielt, umfasst neben fachlichen und sozialen stets auch selbstbezogene Kompetenzen (Reetz, 2010). In interaktionslastigen Dienstleistungsberufen, bei denen psychische Belastungsfaktoren eine besondere Herausforderung darstellen (Böhle & Weihrich, 2020), kommt einer Belastungs-Bewältigungs-Kompetenz (BBK) zur Erhaltung der eigenen Gesundheit und Leistungsfähigkeit ein besonderer Stellenwert als eine Facette selbstbezogener Kompetenzen im beruflichen Handlungskontext zu (Wittmann et al., 2022). Im Gegensatz zu sozialen Kompetenzen (z.B. Kanning, 2015) wurde die BBK bislang aber weder hinreichend theoretisch fundiert noch empirisch systematisch erforscht. Dieses Defizit wiegt insbes. in Gesundheitsberufen wie der Pflege schwer. In diesem Berufsfeld sind die auf die Beschäftigten einwirkenden Belastungsfaktoren quantitativ (z.B. Zeit-Mengen-Problematik angesichts Personalmangels) wie auch qualitativ (z.B. Konfrontation mit Leid und Sterben) virulent und gelten als wesentliche Ursachen für Ausbildungsabbrüche und Kündigungen (Jacobs et al., 2021). Die BBK kann hier also als hoch bedeutsam mit Blick auf eine langfristige Sicherung des Fachpersonals im Gesundheitswesen eingestuft werden (Höhmann et al., 2016). Im Projekt Erweiterte Kompetenzmessung in Gesundheitsberufen (EKGe) wurde ein digital gestütztes Instrument zur Erfassung der BKK angehender Pflegefachkräfte theoriegeleitet entwickelt und curricular wie auch psychometrisch validiert (Warwas et al., im review). Eine Interventionsstudie zur Förderung dieser Kompetenz adressiert folgende Forschungsfragen:

(1) Wie entwickelt sich die BBK im Ausbildungsverlauf und welche Kompetenzniveaus werden am Ausbildungsende erreicht? (diagnostische Perspektive)

(2) Welchen Einfluss haben Arbeits- und Interaktionsmerkmale in den Ausbildungsabschnitten der sog. Praxisanleitung in ausbildenden Pflegeeinrichtungen im Sinne von Stützfaktoren auf die Kompetenzentwicklung? (betriebspädagogische Perspektive).

Frage 2 nimmt sog. Prozessqualitäten der Ausbildung in den Praxiseinrichtungen in den Blick. Sie gründet in der Überlegung, dass der Erwerb von BKK überwiegend informell durch Modelllernen im (teils kooperativen) Arbeitsvollzug erfolgt, aber auch die pädagogische Beziehungsgestaltung durch die Praxisanleitung relevant sein kann (Allmacher & Stähling, 2019). Hinzu kommt die Erkenntnis aus der Stressforschung, dass unterstützende Interaktion im (Arbeits-)Alltag nachweislich zur Belastungsbewältigung beiträgt (Schmucker 2020). Bei den Prozessdimensionen betrieblicher Ausbildungsqualität unterscheiden Böhn und Deutscher (2021) drei Bereiche: (1) Struktur- und Ausführungsmerkmale der zugeteilten Aufgaben und deren Lernpotentiale (z.B. Planungs- und Entscheidungsspielraum) einschließlich Belastungspotentiale (z.B. Komplexität); (2) Merkmale der sozialen Interaktion i.w.S., darunter die Einbindung in die lokale Expertengemeinschaft; (3) Merkmale der pädagogischen Begleitung i.e.S., etwa die mentorielle Betreuung oder individuelles Feedback.

Zur Messung der BKK wurde ein in authentische Handlungssituationen der Pflegeausbildung verankerter Situational Judgement Test aus 9 Situationsvignetten für die drei Ausbildungssettings

eingesetzt. Die Erhebung erfolgte zweimal im Abstand von ca. sechs Monaten. Individuelle Leistungsdaten wurden mithilfe eines Partial-Credit-Modells skaliert. Für die statistischen Analysen stehen Daten von N=247 Auszubildenden in der Pflege zur Verfügung, die die Kompetenztests absolvierten und Fragebögen zu den Ausbildungsbedingungen der zuletzt durchlaufenen Organisationseinheit ihrer ausbildenden Pflegeeinrichtung ausfüllten. Für die längsschnittliche Skalierung der Leistungsdaten wird ein zweidimensionales Modell eingesetzt, das die Items einer zeitlichen Dimension zuordnet. Zur Untersuchung des Erklärungsbeitrags von Stützfaktoren im Ausbildungsprozess zur Entwicklung der BKK werden Strukturgleichungsanalysen eingesetzt, in welche die Skalen Autonomie, Komplexität, Vielfalt zugewiesener Aufgaben, Interaktionserfordernisse bei der Aufgabenerfüllung, Beziehungsqualität und funktionale Einbindung im Team, Feedback und Theorie-Praxis-Verknüpfung durch den/die Praxisanleiter/in als Prädiktoren einfließen.

Die (messzeitpunktspezifische) psychometrische Qualität des 29 Items umfassenden Testinstruments zur Messung der Bewältigungskompetenz wurde mittels Partial-Credit-Modellierung und DIF-Analysen für die pflegerischen Ausbildungsberufe geprüft: die Item-Kennwerte liegen in einem akzeptablen Bereich (0.87 ≤ wMNSQ ≤ 1.14; -1.8 ≤ t-Werte ≤ 2.1; Trennschärfe ≥ 0.26). Zudem sind alle Items für die drei untersuchten Pflegeuntergruppen fair. Das Testinstrument weist eine sehr gute interne Konsistenz auf (EAP/PV Rel. = 0.91). Die Strukturgleichungsanalysen werden gegenwärtig noch durchgeführt. Die Ergebnisse werden mit Blick auf Qualitätsverbesserungen in der Pflegeausbildung diskutiert.

Entwicklung und Prädiktoren der betrieblichen Identifikation von Auszubildenden im ersten Ausbildungsjahr

Elisabeth Maué, Stefanie Findeisen, Stephan Schumann
Universität Konstanz

Theoretischer Hintergrund

Die Herausbildung beruflicher und betrieblicher Identifikation ist ein zentrales Ziel in der beruflichen Ausbildung (Heinrichs et al., 2022). Angenommen wird dabei, dass mit einer höheren Identifikation eine bessere Arbeitsleistung und eine gesteigerte Persistenz einhergeht. Entsprechend sind Unternehmen daran interessiert, die Identifikation der Auszubildenden mit dem Ausbildungsbetrieb von Beginn an zu fördern. Dabei fehlen jedoch bisher ein tieferes Verständnis der Prädiktoren und Prozesse betrieblicher Identifikation sowie empirisch belegte Erkenntnisse zur Identitätsentwicklung Auszubildender (Heinrichs et al., 2022).

Die Identifikationsprozesse Auszubildender sind aufgrund der Bedeutung des Übergangs von der Schule in den Beruf für den beruflichen Erfolg besonders relevant (Dawis & Lofquist, 1984). Gerade zu Ausbildungsbeginn kommen Jugendliche mit der neuen Umwelt „Betrieb“ in Kontakt, weshalb die Studie die betriebliche Identifikation genauer untersucht. Die betriebliche Identifikation beschreibt die psychologische Bindung einer Person an eine bestimmte Organisation (Mael & Ashforth, 1995) und damit eine Form der sozialen Identität (Tajifel & Turner, 1979).

Bisherige Studien zu Auszubildenden zeigen in diesem Zusammenhang berufsspezifische Unterschiede in der Identifikationsausprägung und deren Entwicklung (Heinemann et al., 2009; Metzlaff, 2015). Die Identifikation steht dabei in positivem Zusammenhang mit der Integration am

Arbeitsplatz, der Unterstützung durch Kolleg:innen, dem individuellen Engagement und den beruflichen Kompetenzen (Heinemann et al., 2009; Klotz et al., 2014; Metzlaff, 2015). Während der Berufsausbildung nimmt die betriebliche Identifikation ab, was durch die Erfüllung individueller Grundbedürfnisse (Kompetenz, Verbundenheit, Autonomie; Deci & Ryan, 1993) teilweise kompensiert wird (Kirchknopf & Kögler, 2022). Eine hoch ausgeprägte anfängliche Identifikation und eine schwache Abnahme verringern zugleich die Absicht, den Ausbildungsberuf oder -betrieb zu wechseln (Kirchknopf & Kögler, 2022).

Fragestellungen

1. Wie entwickelt sich die betriebliche Identifikation im ersten Jahr der Berufsausbildung?
2. Welche Faktoren beeinflussen die betriebliche Identifikation?
3. Wie hängen betriebliche Identifikation, Aufgabenbewältigung und Engagement zusammen?
4. Wie wirkt sich die betriebliche Identifikation auf die Zielerreichung aus?

Methode

Wir verwenden Längsschnittdaten einer Studie mit standardisierten Online-Fragebögen, die zu Beginn (t1), nach drei (t2) und nach neun Monaten der Berufsausbildung (t3) eingesetzt wurden. Die Stichprobe umfasst 273 Auszubildende (54% Männer, 46% mit Migrationshintergrund) aus den Bereichen „Handel“ (43%), „Handwerk“ (37%) und „Pflege/Medizin/Erziehung“ (20%). Es werden Instrumente zu sozialer Unterstützung durch Ausbilder:innen und Kolleg:innen (t1), sozialer Integration (t2), formaler Sozialisation (onbarding-Maßnahmen) (t2), betrieblicher Identifikation (t2, t3), Leistungsindikatoren (Zielerreichung und Aufgabenbewältigung; t2, t3) und Engagement (t2, t3) genutzt. Die Fragestellungen werden mittels eines Strukturgleichungsmodells untersucht.

Ergebnisse

Zunächst wird deutlich, dass die Auszubildenden zu t1 viel Unterstützung durch Ausbilder:innen und Kolleg:innen erfahren. Zu t2 berichten sie über ein hohes Maß an Aufgabenbewältigung, Zielerreichung und Engagement. Die betriebliche Identifikation nimmt über die Zeit unwesentlich ab (t2: M=3.44, SD=0.91; t3: M=3.32, SD=0.96; n.s.). Das Engagement sinkt dagegen zwischen t2 und t3 deutlicher. Weitere Analysen zeigen, dass die formale Sozialisation die betriebliche Identifikation zu t2 positiv beeinflusst, welche ihrerseits Engagement und Zielerreichung zu t3 positiv voraussagt. Die Aufgabenbewältigung zu t2 beeinflusst wiederum die betriebliche Identifikation zu t3 und umgekehrt (cross-lagged-Effekt).

Die insgesamt positive Sicht Auszubildender auf die erste Ausbildungszeit ist anschlussfähig an Heinemann et al. (2009) sowie Kirchknopf und Kögler (2022) (F1). Die Ergebnisse unterstreichen die Bedeutung formaler Sozialisation für betriebliche Identifikation (F2; Klotz et al., 2014). Weiterhin zeigen sich positive Effekte betrieblicher Identifikation zu t2 auf Zielerreichung und Engagement zu t3 sowie ein cross-lagged-Effekt zwischen betrieblicher Identifikation und Aufgabenbewältigung (F3&4). Die Ergebnisse bestätigen zugleich die Zusammenhänge zwischen Identifikation und Leistung (Kirchknopf/Kögler, 2022; Klotz et al., 2014) sowie die Bedeutung von Identifikation für Engagement (Heinemann et al., 2009; Metzlaff, 2015). Trotz Limitationen (z.B. selbstberichtete

Leistungsindikatoren) adressiert die Studie Forschungslücken (Heinrichs et al., 2022) und erweitert Erkenntnisse zu Entwicklung und Prädiktoren betrieblicher Identifikation Auszubildender verschiedener Berufsfelder. Die Ergebnisse sind darüber hinaus für die Gestaltung betrieblicher Ausbildung relevant.

F-16: Schulleitung in Zeiten der Krise: Heraus- oder doch Überforderung?

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* T03 R03 D89

Schulleitung in Zeiten der Krise: Heraus- oder doch Überforderung?

Chair(s): **Jana Groß Ophoff** (Pädagogische Hochschule Vorarlberg, Austria)
*Diskutant*in(nen):* **Barbara Herzog-Punzenberger** (Pädagogische Hochschule Wien)

Im deutschsprachigen Raum und darüber hinaus besteht ein anhaltender Mangel an Schulleitungen, die jedoch als wichtige „top-down" Akteure in schulischen Entwicklungsprozessen angesehen werden (Brown et al., 2021). Dies steht vermutlich auch in Verbindung damit, dass die Entscheidung für ein Schulleitungsamt durch ein Streben nach mehr Gestaltungsmöglichkeiten und Verantwortung geprägt ist (Cramer et al., 2021). Angesichts der kontinuierlichen Erweiterung der Anforderungen und Erwartung an Schulleitungen (Cramer & Rothland, 2021; Evans, 1995; Leithwood, 2012) und wiederkehrender Phasen hoher Arbeitsbelastung (Pont et al., 2020) verdeutlichen u.a. Pietsch et al. (2020) die Relevanz, eine Balance zwischen einem Verwalten der Schule mit vorhandenen Mitteln und radikaler Innovation zu finden, um langfristig Überforderung und chronischen Stress zu vermeiden. Die ohnehin angespannte schulische Situation wurde in den letzten drei Jahren durch die Pandemie weiter verschärft, wenn etwa trotz Schulschließungen der Schulbetrieb aufrecht erhalten werden musste. Schulleitungen waren noch mehr als ohnehin als Krisenmanager:innen gefordert (Ansell et al., 2021; Harris & Jones, 2020; Weiner et al., 2021).

Im Symposium werden vier Beiträge mit unterschiedlichem Schwerpunkt auf die Frage eingehen, was Schulleitungshandeln in Krisenzeiten herausfordernd, aber vielleicht auch überfordernd „macht" – und zwar auf Basis von umfangreichen Schulleitungsbefragungen im deutschsprachigen Raum. Im Mittelpunkt des *ersten Beitrags* steht das professionelle Selbstverständnis von deutschen Schulleitungen als „Change Agents", der anhand vertiefter Einblicke den Mehrgewinn durch einen Mixed-Methods-Zugang (Interviews und Surveys) in der empirischen Schulleitungsforschung aufzeigt. Auch der zweite Beitrag fußt auf einem solchen methodischen Vorgehen (Einzel-/Gruppeninterviews und Surveys). Im Vordergrund stehen hier ausgewählte Befunde zur kulturellen Responsitivität von Schulleitungen in Österreich, wonach angesichts zunehmender kultureller Vielfalt wichtig erscheint, in der Schulleitung mit Fragen von Inklusion und Chancengleichheit kritisch-reflexiv umgehen zu können. Der *dritte Beitrag* vergleicht auf Basis von umfangreichen Surveystudien, die in den Jahren 2021 und 2022 durchgeführt wurden, österreichische, schweizerische und deutsche Schulleiter:innen. Erste Ergebnisse für das schweizerische Survey demonstrieren die in der Literatur vielfach als Belastung diskutierte Ambivalenz zwischen sinnstiftenden Tätigkeiten auf der einen Seite und großem Zeitdruck sowie einer als mangelnd erlebten Unterstützung auf der anderen Seite. Auch der *vierte Beitrag* nimmt eine ländervergleichende Perspektive ein, hier aber weiterführend auf das arbeitsbezogene Bewältigungsverhalten sowie Stress und psychische Gesundheit deutscher und schweizerischer Schulleiter:innen.

Die Beiträge werden kritisch vor dem Hintergrund diskutiert, dass es sich bei Schulleitungen um eine für das Bildungssystem wichtige Akteur:innengruppe handelt, die insgesamt noch zu wenig Aufmerksamkeit von bildungspolitischen Entscheidungstragenden zu erhalten scheint, weshalb Eingangsvoraussetzungen, Arbeitsbedingungen, aber auch Unterstützungsmöglichkeiten von Schulleitungen weiter zu erforschen sind.

Beiträge des Symposiums

Die Cornelsen Schulleitungsstudie Die Bedeutung von Mixed-Methods-Ansätzen für eine teilnehmer:innenorientierte Forschung

Martin Bittner[1], Sarah Fichtner[1], Matthias Sandau[2]
[1]FiBS Forschungsinstitut für Bildungs- und Sozialökonomie, [2]m.sandau@fibs.eu

Theoretischer Hintergrund

Obwohl Schulleitungshandeln als eine der zentralen Stellschrauben für ein zukunftsfähiges Bildungssystem zu begreifen ist (Pietsch et al. 2016; Heffernan 2019; Tulowitzki und Pietsch 2020), spielen Schulleitungen eine marginalisierte Rolle in Deutschland, wenn es um die Einbeziehung ihrer Perspektiven in die Analyse und die darauf aufbauenden Impulse für die Entwicklung des Bildungssystems geht. Dies zeigt sich u.a. in einer lediglich angestoßenen Professionalisierungsdebatte (Brändle et al. 2022). Schulleitungen werden hier vor allem in ihrer Wirksamkeit als Führungspersonen hinterfragt (Brauckmann et al. 2020), die weitere gesellschaftliche und politische Anerkennung steht noch aus. Auf institutioneller Ebene entscheiden Schulträger und Schulverwaltung formell über die Ausstattung und Ausrichtung der Schule. Die gelebte Haltung, Praxis und Vision steht und fällt allerdings mit der Schulleitung, die weit mehr ist – und sein muss – als eine mit administrativen Befugnissen ausgestattete Lehrkraft. Aktuelle Befragungen von Schulleitungen zeigen zudem, wie sehr ihr Handeln an einer pädagogischen Zielsetzung der Institution Schule ausgerichtet ist (Fichtner et al. 2022), die sich nicht auf eine fachliche Outputorientierung reduzieren lässt, sondern die Lebenskompetenzen und die Persönlichkeitsentwicklung von Schüler:innen in den Blick nimmt.

Fragestellung

Wie können die Perspektiven von Schulleitungen als Zukunftsgestalter:innen sicht- und hörbar gemacht werden? Wie können wir adäquat transportieren, wie Schulleitungen sich selbst, ihre Aufgaben, ihre Rollen, Führungspraktiken, Bedarfslagen und aktuellen Herausforderungen in Deutschland wahrnehmen? Wie ihre Impulse als Innovator:innen für die Gestaltung der Schule als pädagogische Institution von morgen kommunizieren?

Methode

Die Cornelsen Schulleitungsstudie bringt in einer jährlichen, repräsentativen Umfrage der Schulleitungen allgemeinbildender Schulen in Deutschland die zentralen und brennendsten Bildungsthemen und Entwicklungen auf das Tableau. Dafür befragt sie mittels eines Mixed-Methods-Ansatzes Schulleitungen zu dem, was sie in ihrem Schulleitungsalltag beschäftigt, zu ihrem Führungshandeln, ihren Bewertungen der Kontextbedingungen ihrer Praxis, den von ihnen wahrgenommenen Herausforderungen und Lösungsansätzen, ihren Perspektiven auf Schule, sowie ihren Wünschen und Visionen.

Dies erfolgt in einem ersten Schritt durch ausführliche, leitfadengestützte Interviews. Aus der qualitativ-inhaltlichen Analyse dieser 50 (2021) bzw. 20 (2022) codierten Interviewtranskripte - angereichert durch relevante Literatur - werden Skalen operationalisiert und in einen standardisierten Fragebogen überführt. Im zweiten Schritt wird durch eine repräsentative Erhebung die Verteilung der

Merkmale überprüft, anhand deskriptiver Analysen dargestellt und für einzelne Schwerpunktthemen durch repräsentative Gruppenunterschiede und multivariate Analysen

tiefergehend betrachtet. Die Ergebnisse liefern neben einem repräsentativen Abbild der Stimmen und Stimmungen von Schulleiter:innen einen Ausgangspunkt für künftige Erklärungen von Schulleitungshandeln und durch zukünftige Trendanalysen den Rückschluss auf kontextbedingte Veränderungen des Handelns und der Visionen von Schulleitungen.

Ergebnisse und ihre Bedeutung

Schulleitungen zeichnen sich zumeist als visionäre Gestalter:innen des Bildungssystems aus. Sie erarbeiten und entwickeln eine Schule – präferiert als Team, in Zusammenarbeit mit einem multiprofessionell aufgestellten Kollegium und in Vernetzung mit anderen Schulleitungen –, die sich den aktuellen Bildungsdiskursen zuwendet. Mit verzahnten, individualisierten und kooperativen Lernformaten, im projektorientierten Lernen, mit Kooperationspartnern im Sozialraum, in der Ganztagsschule, mittels sinnvoll eingesetzter, digitaler Tools, sehen Schulleitungen vielfältige Möglichkeiten, die Lernbiografien ihrer Schüler:innen positiv zu beeinflussen.

Ihre Bedarfe sind zugleich politische Themen: eine Ausstattung an qualifiziertem Personal, die es ermöglicht, die heterogenen Anforderungen an die Unternehmung Schule effektiv zu bewältigen; eine digitale Ausstattung und Ausbildung an derselben, die zum Lernen im Kontext von Digitalität befähigt; eine räumlich-bauliche Situation, die flexible Lehr-Lernsettings gestattet. Die Praxis des Schule-Leitens basiert dabei auf wertschätzenden, partizipativen und demokratischen Strukturen, die von allen Beteiligten anerkannt und gelebt werden müssen. Die Schulleitungsstudie liefert nicht nur Erkenntnisse für die Bildungssteuerung, sondern bringt durch das Instrument eines induktiv gestützten Fragebogens neue Inhalte in die Debatten um das Schulleitungshandeln ein.

Kulturelle Responsivität von Schulleitungen in österreichischen Schulen

Sieglinde Fulterer, Alexandra Postlbauer
Linz School of Education, JKU Linz

Theoretischer Rahmen

Die Zunahme der sprachlichen, kulturellen, religiösen und ethnischen Zugehörigkeit der Schüler*innen in den letzten vier Jahrzehnten Bedarf der Etablierung eines kultursensiblen Schulumfeldes (Beachum, 2011; Kahlifa et al., 2016). Insbesondere Schulleitungen sind gefordert, kultursensible Praktiken zu entwickeln, die Diversität einbeziehen, da die pädagogische Praxis der Schulleitung richtungsweisend für die Gewährleistung inklusiver Lernbedingungen und -möglichkeiten aller Schüler*innen sein kann (Dimmock & Walker, 2005; Khalifa, 2018; López, 2016; Magno & Schiff; 2010). Kulturelle Responsivität „could be understood as having a positive, respectful, sensitive reaction to different expressions and habits of culture. In order that all students are welcomed and accepted and inspired to succeed in a culture of high expectations for learning, schools and classrooms must be responsive to culture" (Ontario Ministry of Education, 2013). Dabei wird von einem Kulturbegriff ausgegangen, der nicht monolithisch ist, sondern durch soziale Interaktion (re-)produziert wird, sowie 'Kultur' in den jeweiligen historischen Kontext setzt und Repräsentationsräume ermöglicht, denn „culture not only mediates history but shapes it" (Giroux, 2004, S. 62).

Kultursensibler Dialog wird als Instrument der Verknüpfung von kulturell vielfältigen Perspektiven und des sozialen Zusammenhalts begriffen (Fass et al., 2014, S. 312). Dahingehend ist die soziale Interaktion für pädagogisches Leadership in vielfältigen Communitys relevant (Dimmock & Walker, 2005).

Fragestellung

Basierend auf den Ergebnissen einer Schulleitungsbefragung (N = 148; MAlter = 54.22; SD = 6.22; 58% weiblich, 42% männlich, 0% divers) sowie Einzel- und Gruppeninterviews (Schulleiter*innen, Lehrpersonen, Erziehungsberechtigte, Schüler*innen; N = 61; 69% weiblich, 31% männlich, 0% divers; 44% mit Migrationshintergrund; 0% mit Fluchthintergrund), die im Rahmen des dreijährigen Erasmus+ Projekts Supporting Culturally Responsive Leadership and Evaluation in Schools (CReLES) durchgeführt wurden, werden (1.) die Prädiktoren kultureller Responsivität von Schulleitungen sowie (2.) die schulischen Praktiken als Reaktion auf die zunehmende kulturelle Vielfalt und die Bedingungen und bewährten Strategien für eine kultursensible Schulentwicklung, um bessere Lernergebnisse für alle Schüler*innen zu gewährleisten, näher beleuchtet.

Methoden

Mittels Regressionsanalysen wurde der Zusammenhang zwischen (individuellen) Merkmalen der Schulleitung, der Schule, der elterlichen Inklusion, des Stakeholder-Supports, der wahrgenommenen Herausforderungen und der kulturresponsiven Schulpraxis eruiert (OECD, 2018; Schulpolitik und Unterrichtsstrategien: 7 Items, α = .84; Zusätzliche Unterstützung für Schüler*innen: 3 Items, α = .72; Intendierte Lernerfahrungen: 4 Items; α = .91; Lernerfahrungen außerhalb des Curriculums: 4 Items, α = .67). Die Auswertung der Interviews der Fallschulen wurde mittels qualitativer Inhaltsanalyse nach Mayring (2014) auf Grundlage von Kategorien für die deduktive Kodierung und zusätzlicher induktiv generierter Codes vollzogen. Zu berücksichtigen ist, dass aufgrund der Auswahl von Best-Practice-Fallschulen nicht davon ausgegangen werden kann, dass die Ergebnisse verallgemeinert werden können. Anhand der Analyse können jedoch Phänomene und Zusammenhänge in kultursensiblen Schulentwicklungsprozessen identifiziert und nachvollzogen werden, die in einzelnen Schulen beobachtbar sind.

Ergebnisse und ihre Bedeutung

Deskriptive Analysen verdeutlichen die Bereitschaft der Schulleiter*innen zu kultureller Responsivität (5-stufige Skala; M = 3.10; SD = .63). Individuelle Merkmale der Schulleitung stehen nicht in Zusammenhang mit kultureller Responsivität, wohingegen Merkmale des Schulstandorts (z.B. Gemeindegröße: $R2 = .20$, $\beta = .45$, $p \leq .001$), die Zusammensetzung der Schüler*innenschaft ($R2 = .19$, $\beta = .40$, $p \leq .05$) wie auch auch kontextspezifische Merkmale (z.B. Inklusion der Eltern $R2 = .10$, $\beta = .27$, $p \leq .01$) kultursensible Praxis beeinflussen. Die qualitativen Inhaltsanalysen der Interviews hingegen zeigen, dass die Erfahrungen, Einstellungen und Kompetenzen der Schulleitung wichtige Faktoren für kultursensible Praxis an Schulen sind. Kulturelle und sprachliche Vielfalt werden als Chance für alle Schüler*innen interpretiert, während gleichzeitig auf die Notwendigkeit des Erlernens der Unterrichtssprache hingewiesen wird. Bestimmte kultursensible Praktiken, wie die kritische

Überprüfung der Schulpolitik und -praxis im Hinblick auf Inklusion und Chancengleichheit, erscheinen dabei als wesentlich für die Verbesserung der kultursensiblen Praxis.

Quo vadis Schulleitung? Befunde aus trinationalen Erhebungen zu Karrieremotiven, zu Arbeitsplatzwechselabsichten und zum Führungsverständnis

Pierre Tulowitzki[1], Marcus Pietsch[2], Colin Cramer[3], Jana Groß Ophoff[4], Ella Grigoleit[1], Gloria Sposato[1]
[1]Fachhochschule Nordwestschweiz, [2]Leuphana Universität, Lüneburg, [3]Universität Tübingen, [4]Pädagogische Hochschule Vorarlberg

Theoretischer Hintergrund

Schulleitungen und ihr Handeln werden als relevante Grösse lernwirksamer Schulen (Day et al., 2016) und erfolgreicher Schulentwicklungsprozesse (Buchen et al., 2009) angesehen. Das schulische Leitungshandeln wirkt sich maßgeblich auf die strukturelle Gestaltung innerhalb der Organisation, das Schulklima und die Schulkultur sowie die Zufriedenheit, Motivation und Leis-tungsfähigkeit des schulischen Personals aus (u.a. Day et al., 2009; Seashore Louis et al., 2010). In Deutschland, Österreich und in der Schweiz liegt – unterschiedlich ausgeprägt – ein Mangel an Schulleitungen vor. Für Deutschland gibt es zudem Hinweise, dass Schulleitungs-wechsel eher häufig vorkommen (Pietsch et al., 2019). Um Personen für das Amt gewinnen und in ihm halten zu können, erscheinen Forschungsbefunde dazu wichtig, was das Amt der Schul-leitung attraktiv macht, warum Schulleitungen diese Position ergreifen und welche Gründe gegen einen Verbleib im Amt sprechen.

Fragestellungen

Im Zentrum der vorliegenden Studie standen folgende Fragen: Warum ergreifen Schulleitungen diese Tätigkeit? Was (de-)motiviert sie in ihrem Beruf? Wie charakterisieren sie ihre berufliche Identität und ihr Führungshandeln?

Methode

Die vorliegende Studie wurde 2021 in der Schweiz und wird im Herbst 2022 auch in Deutschland und Österreich durchgeführt. In der Schweiz wurden die Daten im Herbst 2021 mit Unterstützung aller Schulleitungsverbände und zahlreicher Kantone mittels eines standardisierten Online-Fragebogens erhoben, der an 4749 Schulleitungen verschickt wurde (ca. 52% aller Schulleitun-gen). 2011 Schulleitende beteiligten sich vollständig an der Befragung. In Weiterführung früherer Befragung wurden in Deutschland im Oktober 2022 1000 Schulleitungen durch den deutschen Felddienstleister forsa rekrutiert und befragt. Teile der Stichprobe wurden bereits seit 2020 be-fragt (Cramer et al., 2021; Pietsch et al., 2022), so dass z.T. längsschnittliche Analysen möglich sind. In Österreich wird die Studie im Auftrag durch das österreichische Bildungsministerium durch die PH Vorarlberg im Herbst 2022 erstmalig durchgeführt und alle österreichischen Schul-leitungen dazu eingeladen (ca. 6000 Schulen).

Die Fragebögen bestehen überwiegend aus etablierten, thematisch relevanten Skalen (u.a. aus PISA, TIMSS, TALIS) und Items zur Erfassung demografischer Angaben. Um einem Common Method Bias entgegen zu wirken, wurden Cooper et al. (2020) folgend, wo möglich Items und Skalen rotiert und unterschiedliche Skalierungsgrade genutzt. Die deskriptiven Statistiken werden mit der Software SPSS 26 erstellt; multivariate Analysen stehen für die Schweiz noch aus. Erfasst wurden bzw. werden

u.a. folgende Konstrukte: Berufliche Werthaltungen, berufliches Selbstverständnis, Arbeitszufriedenheit, Karriereperspektiven, Füh-rung.

Ergebnisse

Der Ergebnisteil des vorliegenden Beitrags rekurriert auf die Befunde aus der Schweiz; zum Zeitpunkt der GEBF werden jedoch auch die Befunde aus Österreich, aktualisierte Befunde aus Deutschland sowie erste Vergleiche zwischen den drei Ländern vorliegen.

Oft genannte Gründe, in die Schulleitung zu gehen, sind gemäß ersten Analysen des schweize-rischen Datensatzes: Verantwortung übernehmen und für das Wohl anderer sorgen wollen; Geld und Karrieremotive werden vergleichsweise selten genannt. Das rückgemeldete Tätigkeitsprofil besteht überwiegend aus administrativen und personalbezogenen Aufgaben. Ein Großteil der Schulleitenden äußert sich weitgehend zufrieden mit der beruflichen Karriere und erlebt die Tä-tigkeit als sinnstiftend. Jedoch geben 74% der Befragten an, trotz Mehrarbeit zu wenig Zeit zu haben, die täglichen Aufgaben zu erfüllen. Ca. 20% der Befragten geben an, ihre Schule verlas-sen zu wollen; ein Viertel von ihnen so schnell wie möglich. Hauptsächliche Gründe hierfür sind der Wunsch nach beruflicher Entwicklung und nach mehr Zeit (für die Familie) sowie fehlende Unterstützung durch übergeordnete Instanzen.

Diskussion

Die Befragten nehmen sich mehrheitlich eher als professionelle Führungskräfte denn als Admi-nistrator:innen oder Lehrpersonen mit Zusatzfunktion war und sehen sich zuständig für Entwick-lungsimpulse. Die Aussage, dass Leistung der Schulleitungen nach den Lernergebnissen der Schüler:innen zu bemessen sei, wird überwiegend abgelehnt. Die Diskussion wird abschließend u.a. Gründe für etwaige Unterschiede zwischen den drei Ländern im Schulleitungshandeln und in den auf das Amt bezogenen Einstellungen und Haltungen aufgreifen und insgesamt die Befunde in den Stand der Forschung einordnen.

Arbeitsbezogener Stress, Bewältigungsverhalten und emotionale Erschöpfung von Schulleitungen aus Deutschland und der Schweiz. Ergebnisse der COVID-HL Studie

Kevin Dadaczynski[1], Anita Sandmeier[2], Orkan Okan[3], Melanie Messer[4]
[1]Fachbereich Gesundheitswissenschaften, Hochschule Fulda, Zentrum für Angewandte Gesundheitswissenschaften, Leuphana Universität Lüneburg, Deutschland, [2]Institut Professionsforschung und Personalentwicklung Pädagogische Hochschule Schwyz, Schweiz, [3]Fakultät für Sport- und Gesundheitswissenschaften, Technische Universität München, Deutschland, [4]Abteilung Pflegewissenschaft II, Universität Trier, Trier, Deutschland

Hintergrund: Während sich Schüler*innen sowie Lehrkräfte im Rahmen der empirischen Schulgesundheitsforschung eines hohen Interesses erfreuen, steht die Forschung zur Gesundheit von Schulleitungen bislang noch an den Anfängen. Die wenigen verfügbaren Befunde weisen u. a. darauf hin, dass Schulleitungen häufig von psychischen Belastungen und Beanspruchungen betroffen sind, die sich im Berufsgruppenvergleich oder dem Vergleich einer bevölkerungsbezogenen Normgruppe eine höhere Ausprägung aufweisen (Dewa et al., 2009; Philips et al., 2008). Ergebnisse im Kontext der seit 2020 in Europa bestehenden COVID-19 Pandemie zeigen, dass sich Schulleitungen vor allem durch die Arbeitsmenge, interpersonale Konflikte, einen Mangel an Ressourcen und persönlichen Druck gestresst fühlen (Hansen et al., 2020). Mit Blick auf die physische und psychische Erschöpfung geben 45 % der in einer deutschen Studie befragten Schulleitungen und Schulleitungsmitglieder an, dass es ihnen am Ende eines Arbeitstages

schwerfällt, sich zu erholen und ein Drittel fühlt sich oft/immer psychisch erschöpft (Dadaczynski et al., 2021). In einer Schweizer Studie berichteten 29 % der Schulleitungen ein reduziertes Wohlbefinden, einem Viertel fällt es schwer, sich nach einem Arbeitstag zu erholen und jede fünfte Schulleitungsperson fühlt sich bei der Arbeit oft psychisch erschöpft (Sandmeier et al., 2021).

Wie Schulleitungen mit arbeitsbezogenem Stress umgehen und inwiefern arbeitsbezogenes Bewältigungsverhalten mit der psychischen Gesundheit assoziiert ist, ist bislang nicht untersucht worden.

Fragestellung: Ausgehend vom Job Demands-Resources Modell (Demerouti et al., 2001) und der transaktionalen Stresstheorie (Lazarus & Folkman, 1984) wird der Zusammenhang zwischen arbeitsbezogenem Stress und psychischer Gesundheit von Schulleitungen während der COVID-19 Pandemie untersucht. Überdies soll geklärt werden, wie Schulleitungen arbeitsbezogene Belastungen bewältigen und inwiefern das Bewältigungsverhalten den Zusammenhang zwischen Stress und psychischer Gesundheit erklären kann.

Methoden: Es werden Daten der COVID-HL Schulleitungsstudie herangezogen, die in der ersten Jahreshälfte 2021 mit N=2186 Schulleitungen aus Deutschland und N=385 Schulleitungen aus der Schweiz durchgeführt wurde. Zur Erfassung der psychischen Gesundheit wird auf die Erschöpfung als Leitsymptom des Burnouts zurückgegriffen, welche über drei Items und einer fünfstufigen Antwortskala erfasst wurde (Schaufeli et al., 2020). Arbeitsbezogener Stress wurde mithilfe einer Kurzversion der Perceived Stress Scale gemessen, die sprachlich an die COVID-19 Pandemie und Arbeitskontext Schule angepasst wurden (Schneider et al., 2020). Um erste Erkenntnisse zum arbeitsbezogenen Bewältigungsverhalten zu erhalten, wurde auf drei Subskalen (Intensivierung,

Extensivierung, Qualitätsreduktion) des Instruments zur Erfassung des selbstgefährdenden Arbeitsverhaltens zurückgegriffen (Krause et al., 2015). Als Kontrollvariablen werden neben dem Geschlecht und dem Alter ebenfalls die Schulform berücksichtigt. Mittels Strukturgleichungsmodellen in R (R Team, 2020) unter Verwendung der integrierten Entwicklungsumgebung und der Software RStudio mit dem Paket lavaan (Rosseel, 2012) wird das Zusammenwirkens von Stress, Bewältigungsverhalten und Erschöpfung modelliert.

Ergebnisse und ihre Bedeutung: Die Daten befinden sich aktuell in der Auswertung, wobei davon ausgegangen wird, dass arbeitsbezogener Stress positiv mit dem Ausmaß an Erschöpfung assoziiert ist. Ebenso wird davon ausgegangen, dass alle Dimensionen der Arbeitsbewältigung (Intensivierung, Extensivierung und Qualitätsreduktion) mit einer höheren Ausprägung der Erschöpfung einhergehen und dass der Zusammenhang von arbeitsbezogenem Stress und Erschöpfung mindestens partiell durch die Arbeitsbewältigung erklärt wird. Unter Berücksichtigung von Geschlecht, Alter und Schulform sollen die Ergebnisse einen Beitrag zur zielgruppenspezifischen Gesundheitsförderung einer bislang vernachlässigten Berufsgruppe leisten.

F-17: Musikalische Entwicklung im Kindes- und Jugendalter im Zusammenspiel mit kognitiven, motivationalen und psychosozialen Faktoren

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* T03 R04 D10

Musikalische Entwicklung im Kindes- und Jugendalter im Zusammenspiel mit kognitiven, motivationalen und psychosozialen Faktoren

Chair(s): **Patrizia Bieber** (Universität Tübingen, Deutschland), **Richard Göllner** (Universität Tübingen, Deutschland)
*Diskutant*in(nen):* **Henriette Gärtner** (Staatliche Hochschule für Musik Trossingen)

Musik und dem Musizieren werden gemeinhin eine Reihe von positiven Effekten zugesprochen. Dabei werden vor allem Zusammenhänge mit Persönlichkeitsmerkmalen sowie kognitiven Fähigkeiten angenommen. Gleichzeitig ist davon auszugehen, dass auch musikalische Entwicklungsprozesse durch eine Reihe individueller sowie kontextueller Faktoren bestimmt werden. Dies schließt sowohl kognitive und motivationale Eigenschaften auf Seiten von Heranwachsenden als auch auf Seiten ihres familiären Umfeldes mit ein. Entsprechend ergibt sich ein Geflecht wechselseitiger Einflussnahme, dessen empirische Prüfung allerdings nach wie vor aussteht.

Viele Fragen rund um die persönlichkeitsfördernde Wirkung von Musik und ihre Zusammenhänge mit kognitiven, motivationalen und psychosozialen Faktoren sind weithin ungeklärt (Gembris, 2003; Gembris et al., 2015). Studien wie die Langzeitstudie von Bastian (2000) wurden mehrfach hinterfragt und seit Entdeckung des vermeintlichen „Mozart-Effekts“ (Rauscher et al., 1993) sind Musik und ihre Zusammenhänge mit menschlichem Lernen vermehrt im Fokus von Neurowissenschaften und kognitiver Psychologie (Jäncke, 2008; Jentschke & Koelsch, 2006). Befunde deuten darauf hin, dass musikalische Aktivität Spuren in unserem Gehirn hinterlässt und aufgrund dessen Plastizität Einfluss auf die menschliche Entwicklung zu nehmen vermag. Allerdings besitzen existierende Studien im Hinblick auf Wechselwirkungen mit kognitiven, motivationalen und psychosozialen Faktoren aufgrund ihrer querschnittlichen Forschungsdesigns häufig nur eine eingeschränkte Aussagekraft oder liefern ein widersprüchliches Bild (Oliveira et al., 2021; Gembris et al., 2015). Zudem mangelt es nach wie vor an Evidenz zu längerfristigen Effekten von Musik und dem Musizieren auf der Grundlage aussagekräftiger Daten.

Auf unterschiedlichen Längsschnitt-Daten beruhend, betrachten die Symposiumsbeiträge die Thematik aus vier Blickwinkeln.

Beitrag 1 basiert auf Daten, die im Rahmen der Langzeitstudie LongGold erhoben wurden. In den Blick genommen werden subjektive Theorien über Musikalität und deren Einfluss auf musikalische Aktivitäten von Jugendlichen.

Beitrag 2 wechselt die Perspektive und blickt von Elternseite auf den Erwerb musikalisch-spieltechnischer Fertigkeiten. Die Längsschnittstudie ELEMUG geht der Frage nach, inwiefern sich elterliche Überzeugungen im Kontext Instrumentalunterricht über die Zeit verändern und das Lernverhalten von Kindern am Instrument beeinflussen.

Beitrag 3 beleuchtet Zusammenhanganalysen zwischen musikalischer Teilhabe und Persönlichkeit und nutzt dazu Längsschnitt-Daten aus dem Projekt „Musik und Persönlichkeit“. Aufgezeigt werden Interaktionen zwischen musikalischer Teilhabe und den fünf Big-Five-Persönlichkeitsdimensionen.

Beitrag 4 präsentiert vorläufige Ergebnisse der Langzeitstudie LongGold. Der Beitrag beschäftigt sich mit der Fragestellung, inwiefern die Entwicklungen musikalischer und kognitiver Fähigkeiten miteinander in Verbindung stehen, und betrachtet die Bedeutsamkeit gegenwärtiger musikalischer Aktivität als Einflussfaktor.

Diskutiert werden die Beiträge aus biographischer Sicht einer Konzertpianistin.

Beiträge des Symposiums

Einfluss Subjektiver Theorien über Musikalität auf die Entwicklung der Musikalischen Aktivität von Schüler*innen

Miriam Eisinger
Pädagogische Hochschule Freiburg

Was bewegt Kinder und Jugendliche zum Umgang mit Musik? Diese Frage beschäftigt verschiedene Forschungsdisziplinen, wobei unterschiedliche Determinanten kultureller Partizipation diskutiert werden (Kröner, 2013). Auch die zunehmende Anzahl der Programme zur Förderung des Umgangs mit Musik innerhalb und außerhalb der Schule zeigt die aktuelle Relevanz dieser Frage (Krupp-Schleußner, 2018). Dabei stellt die Erforschung des Einflusses der Subjektiven Theorien auf die musikalische Aktivität als Teil kultureller Partizipation ein Desiderat dar. Die Überzeugung, dass sich persönliche Eigenschaften wie z. B. Musikalität verändern lassen und formbar sind, zeigt sich jedoch gerade in herausfordernden Situationen und bei Misserfolgen als wesentlicher Prädiktor für Anstrengungsbereitschaft und Ausdauer (Dweck, 2000).

Das Ziel der Studie ist deshalb, den Einfluss einer dynamischen bzw. stabilen Subjektiven Theorie über Musikalität auf die Entwicklung der musikalischen Aktivität bei Jugendlichen zu untersuchen.

Die Stichprobe wurde im Rahmen des LongGold-Projekts in den Jahren 2018 bis 2022 erhoben und besteht aus 1821 Schüler*innen (51.04% weiblich; Alter zu Studienbeginn: M=10.49 Jahre, SD=0.58; Altersspanne: 9 bis 13 Jahre; Zeitspanne zwischen Testungen: M=362.82 Tage) aus acht Schulen in Niedersachsen, Hessen und Baden-Württemberg, die alle an mindestens zwei von vier Messzeitpunkten teilgenommen haben. Die Daten wurden im Kohorten-Sequenz-Design erhoben, was den möglichen Einfluss von Corona-Maßnahmen gering halten sollte, jedoch noch zu überprüfen ist. Über alle Messzeitpunkte hinweg wurden Subjektive Theorien über Musikalität, d. h. ob Musikalität als etwas Dynamisches oder Stabiles wahrgenommen wird (Eisinger, 2021), und die gegenwärtige musikalische Aktivität (Müllensiefen et al., 2015) erfasst. Die Datenanalysen beinhalten Regressionsanalysen für die Querschnittsdaten und Mehrebenenanalysen für die Längsschnittdaten, wobei die verschiedenen Messwerte innerhalb der Schüler*innen geschachtelt sind (R-Pakete stats und lme4).

Die Ergebnisse der Regressionsanalysen zeigen, dass eine stabile Subjektive Theorie über Musikalität einen negativen (β=-.31, p<.001) und eine dynamische Subjektive Theorie über Musikalität einen positiven (β=.49, p<.001) Einfluss auf die musikalische Aktivität hat. Die Ergebnisse der Mehrebenenanalysen zeigen, dass die individuelle musikalische Aktivität bei allen Schüler*innen über die Zeit hinweg abnimmt (β=-.09, p<.001; BIC=15505). Jedoch legt das Modell, das die Interaktion zwischen Zeit und stabiler Subjektiver Theorie über Musikalität miteinbezieht, aufgrund des besseren Modellfits nahe, dass insbesondere bei den Schüler*innen, die eine stark ausgeprägte stabile

Subjektive Theorie aufweisen, die musikalische Aktivität über die Zeit hinweg stärker abnimmt (β=-.04, p<.001; BIC=15441) als bei denen mit einer dynamischen.

Unsere Studie zeigt, dass eine dynamische Subjektive Theorie über Musikalität ein pädagogisch sinnvoller Ansatz ist und gefördert werden sollte. Im schulischen Musikunterricht und außerschulischen Instrumentalunterricht sollten Jugendliche immer wieder die Erfahrung machen, dass sie etwas lernen und sich weiterentwickeln können. Eine wesentliche Rolle nimmt dabei das Feedback der Lehrperson ein, das nicht die Eigenschaft (z. B. Musikalität), sondern die unternommenen Anstrengungen der Jugendlichen hervorheben sollte, um dynamische Überzeugungen im Unterricht zu fördern (Adams, 2019).

Elterliche Überzeugungen und das Erlernen eines Instruments im Grundschulalter

Patrizia Bieber[1], Barbara Busch[2], Jessika Golle[1], Richard Göllner[1]
[1]Universität Tübingen, [2]Staatliche Hochschule für Musik und Darstellende Kunst Mannheim

Überzeugungen von Eltern beeinflussen das Lernverhalten ihrer Kinder und sind daher unmittelbar an deren Lern- und Entwicklungsprozessen beteiligt. Bisherige Studien zeigen, dass elterliche Überzeugungen (Fan & Chen, 2001; Simpkins et al., 2012) eng mit dem schulischen Erfolg von Kindern verknüpft sind. Sie wirken sowohl auf die persönlichen Erfolgserwartungen des Kindes (Thomas & Strunk, 2017; Simpkins et al., 2012), dessen Wertschätzung des Lerngegenstands (Acosta & Hsu, 2014; DeWitt et al., 2013) als auch auf dessen Selbstbewertung (Parsons et al., 1982; Eccles, 1993), was sich wiederum in der Motivation und den schulischen Leistungen niederschlägt. Gleichzeitig wurden elterliche Überzeugungen häufig nur in Bezug auf die klassischen Lern- und Leistungsdomänen untersucht (z. B. Mathematik; Frenzel et al., 2010; Šimunović & Babarović, 2020), während der musikalisch-künstlerische Bereich vergleichsweise wenig Beachtung fand.

Bislang vorhandene Studien im Kontext Instrumentalunterricht sprechen durchaus dafür, dass elterliche Überzeugungen auch beim Erlernen eines Instruments von Bedeutung sind. Allerdings haben diese Studien lediglich einzelne Facetten elterlicher Überzeugungen beleuchtet (Dai & Schader, 2001; Comeau et al., 2015; Fredricks et al., 2005) oder vielmehr Erwartungen und Ziele in den Blick genommen (Youm, 2008; Egilmez & Engur, 2017; McPherson & Davidson, 2002). Dabei deuten die wenigen existierenden Befunde darauf hin, dass elterliche Überzeugungen nicht unerheblich an spieltechnisch-musikalischen Lernprozessen von Kindern beteiligt sind und aufgrund ihrer Veränderbarkeit besondere Relevanz für die Praxis aufweisen können.

Die vorliegende Arbeit möchte Überzeugungen von Eltern daher aus empirisch-musikpädagogischer Perspektive und in Bezug auf das Erlernen eines Instruments im Grundschulalter in den Blick nehmen. Hierfür wurde im Rahmen des Projektes „Der Einfluss elterlicher Einstellungen und Überzeugungen auf den Erwerb spieltechnisch-musikalischer Fertigkeiten bei Grundschulkindern“ (ELEMUG) ein mehrdimensionales Instrument zur Erfassung elterlicher Überzeugungen im Kontext Instrumentalunterricht entwickeln, um Zusammenhänge mit der Motivation und dem Lernfortschritt von Kindern am Instrument differenziert zu untersuchen.

Konkret wurde untersucht

a) welche Überzeugungen bei Eltern in Hinblick auf das Musizieren und das Erlernen eines Instruments bestehen und inwieweit diese mit der Motivation ihrer Kinder assoziiert sind;

b) inwieweit sich elterliche Überzeugungen auf die Übesituation daheim auswirken und einen Einfluss auf das kindliche Übeverhalten haben und

c) inwieweit elterliche Überzeugungen das Lernverhalten von Kindern im Instrumentalunterricht beeinflussen und damit auf den Lernerfolg am Instrument einwirken.

Die Bearbeitung der drei Fragestellungen erfolgt im Rahmen einer Längsschnittstudie, welche drei Online-Datenerhebungen an 25 Musikschulen in Baden-Württemberg einschließt und im November 2021 startete. Befragt wurden N = 176 Grundschulkinder, die neu mit dem Instrumentalunterricht begonnen haben, deren Eltern (N = 162) sowie deren Lehrkräfte (N = 73). Die vorliegenden Ergebnisse basieren auf Datenanalysen der ersten zwei Messzeitpunkte.

Der entwickelte Elternfragebogen umfasst 51 Items, die sich entlang einer 12-dimensionalen Struktur empirisch trennen ließen. Externe Korrelate der Skalen fanden sich mit dem sozioökonomischen Hintergrund der Eltern, ihrer musikalischen Bildungsbiografie sowie der Lernmotivation ihrer Kinder. Insbesondere die Überzeugungen von Eltern zur Bedeutsamkeit des Musizierens für die Persönlichkeitsentwicklung sowie zur Eigenverantwortung ihres Kindes erwiesen sich dabei als bedeutsam. Kinder, deren Eltern der Überzeugung sind, dass Musizieren eine wichtige persönlichkeitsbildende Funktion erfüllt, berichteten eine höhere Motivation und Investitionsbereitschaft. Ebenso war eine höhere kindliche Motivation mit der elterlichen Überzeugung assoziiert, Kinder besäßen ein hohes Maß an Eigenverantwortung in Bezug auf ihre musikalische Ausbildung.

Ergebnisse des längsschnittlichen Vergleichs darauf hin, dass die elterliche Überzeugung hinsichtlich der Bedeutsamkeit von Fleiß für den instrumentalen Lernfortschritt über die Zeit zunimmt, die elterliche Überzeugung, das Üben kontrollieren zu müssen, dagegen nachlässt. Dabei zeigten sich Unterschiede bei Eltern mit und ohne Erfahrung im Instrumentalspiel, mit und ohne Migrationshintergrund sowie Gruppen unterschiedlicher beruflicher Ausbildung.

Die Befunde bestätigen den Stellenwert elterlicher Überzeugungen hinsichtlich der musikalischen Ausbildung von Kindern bzw. ihre einflussnehmende Rolle beim Erlernen eines Instruments im Grundschulalter.

Mythen kultureller Bildung: Zusammenhangsanalysen von musikalischer Teilhabe und Persönlichkeit im Längsschnitt

Johannes Hasselhorn[1], Valerie Krupp[2]
[1]Hochschule für Musik und Darstellende Kunst Frankfurt, [2]Hochschule für Musik Mainz

„Kulturelle Bildung ist für die Persönlichkeitsentwicklung junger Menschen unverzichtbar" (KMK, 2013, S. 2). So wird auch mit Blick auf musikalische Bildungsangebote deren positive Auswirkung auf die Persönlichkeitsentwicklung junger Menschen vielfach postuliert (z.B. Gembris, 2015, S. 2). Gleichzeitig finden sich bislang für diese Annahme kaum empirischen Belege, zumal die Richtung der kausalen Zusammenhänge in den üblichen Querschnittsuntersuchungen nicht eindeutig geklärt werden kann (vgl. Hasselhorn, et al., 2022). So gibt es zwar Forschung zur typischen Musikerpersönlichkeit (Kemp, 2005), die im Sinne von Expertiseforschung möglicherweise als

Ergebnis weit überdurchschnittlich intensiver Beschäftigung mit Musik angesehen werden können, für eine potentielle Breitenwirkung von kultureller Praxis auf Persönlichkeitsentwicklung im Jugendalter existiert bislang allerdings lediglich anekdotische Evidenz, die aus wissenschaftlicher Sicht den Status eines “Mythos” (Rat für Kulturelle Bildung, 2013) noch nicht überwunden hat. Bisherige Forschungsergebnisse legen umgekehrt nahe, dass individuelle Persönlichkeitsausprägungen kulturelle Teilhabe determinieren (z.B. Kröner et al., 2008; Kröner, 2013). Aufgrund der anhaltenden Erfahrungsberichte aus der Praxis über persönlichkeitsbildende Effekte musikalisch-kultureller Teilhabe beschäftigt sich dieser Beitrag daher mit der Frage, wie Teilhabe und Persönlichkeitsmerkmale sich über längere Zeit gegenseitig beeinflussen.

Zur Beantwortung der Forschungsfrage werden Daten aus dem Projekt “Musik und Persönlichkeit” (2018-2022) herangezogen. Für die Analysen liegen Daten von 1.385 Schüler:innen aus den Klassenstufen 5 bis 7 zu drei Messzeitpunkten vor. Die Schüler:innen wurden im Abstand von jeweils etwa einem Jahr mithilfe digitaler Fragebögen während der Unterrichtszeit befragt. Persönlichkeitsmerkmale wurden mithilfe des Big-Five-Inventars (Kupper et al., 2019) erhoben. Teilhabe an Musikkultur wurde in zwei Dimensionen (formal und informell) erfasst. Zusätzlich wurden demographische und sozio-kulturelle Hintergrundvariablen erfragt. Die statistischen Auswertungen erfolgten mittels Cross-lagged-panel Auswertungen in MPlus7.

Die Ergebnisse zeigen, dass musikalische Teilhabe mit vier der fünf Big-Five-Persönlichkeitsdimensionen interagieren. Für Neurotizismus, Offenheit, Verträglichkeit und Gewissenhaftigkeit können empirische Hinweise dafür gefunden werden, dass Persönlichkeitsausprägungen zukünftige musikalische Teilhabe determiniert, für Verträglichkeit und Gewissenhaftigkeit zeigt sich ebenfalls der umgekehrte Effekt, dass musikalische Teilhabe systematisch die Persönlichkeitsentwicklung beeinflusst. Dabei zeigen sich Unterschiede zwischen formaler und informeller Teilhabe.

Die Ergebnisse werden hinsichtlich möglicher Implikationen für die pädagogische Praxis sowie für weitere empirische Forschung diskutiert.

Die Entwicklung kognitiver und musikalischer Fähigkeiten: vorläufige Ergebnisse der Langzeitstudie Longgold

Daniel Müllensiefen
Goldsmiths, University of London

Längsschnittstudien eröffnen eine differenzierte Perspektive auf die Entwicklung musikalischer Fähigkeiten. Potentiell bieten Längsschnittstudien auch die Möglichkeit, Hinweise auf mögliche kausale Mechanismen und langfristige Konsequenzen musikalischen Trainings zu geben, die in Querschnitts- oder Interventionsstudien nicht oder nur schwer identifiziert werden können. Der vorliegende Beitrag stellt vorläufige Ergebnisse der Längsschnittstudie LongGold mit einem Fokus auf Fragen zur Entwicklung musikalischer und kognitiver Fähigkeiten und der Rolle musikalischer Ausbildung vor.

Der hier verwendete Datensatz umfasst Daten von 4.333 Schülerinnen und Schülern (SuS) an insgesamt 13 weiterührenden Schulen aus Deutschland und England. Das mittlere Alter der SuS betrug 11,8 Jahre bei Eintritt in die Studie (SD = 1,7; Spanne 9-17). 57,2% der Stichprobe gab

‚weiblich' als Geschlecht an und die mittlere Anzahl der Studienteilnahmen lag bei 2,6 (SD= 2,6). Die nachfolgenden Analysen und Modelle basieren auf drei Tests musikalischer Wahrnehmungsfähigkeiten, dem Melodic Discrimination Test (Harrison et al., 2017), dem Beat Alignment Test (Harrison & Müllensiefen, 2018), dem Mistuning Perception Test (Larrouy-Maestry et al., 2019), sowie einem Test zur Kapazität des Arbeitsgedächtnisses (Tsigeman et al., 2022) und einem Test zur fluiden Intelligenz (Chan & Kosinski, 2015). Zusätzlich wurden die gegenwärtigen musikalischen Aktivitäten der SuS erfasst (Müllensiefen et al., 2015) und die musikalische Ausbildung als Facette der musikalischen Erfahrenheit (Müllensiefen et al., 2014). Die Messwerte der drei Musiktests wurden zusätzlich zu einem Faktor Allgemeine musikalische Fähigkeit aggregiert (Pausch et al., 2022).

Die Modellierung der Daten folgte dem Analyseansatz zu Entwicklungsdaten nach McArdle & Nesselroade (2014), der in fünf Schritte gegliedert ist. Im ersten Schritt zeigt sich, dass sowohl die betrachteten musikalischen wie auch die kognitiven Fähigkeiten mit Raten zwischen 11% und 26% einer Standardabweichung pro Jahr wachsen. Das größte Wachstum zeigen die allgemeine Intelligenz und die Beat-Wahrnehmungsfähigkeit. Die zweite Analyse zeigt, dass individuelle Unterschiede in den Wachstumskurven auf allen Fähigkeitsvariablen mit dem Ausmaß an musikalischer Ausbildung positiv assoziiert sind. Die Wachstumsrate für Intelligenz ist jedoch signifikant höher als die Wachstumsrate für allgemeine musikalische Fähigkeit, wie die dritte Analyse zeigt. Im vierten Schritt erweist sich gegenwärtige musikalische Aktivität (CCM) als signifikanter Katalysator für die Entwicklung musikalischer und kognitiven Fähigkeiten. Die Interaktion mit CCM hat eine Effektstärke zwischen 0,04 und 0,12 $\Delta R2$, d.h. sie erklärt zusätzlich 4-12% der Varianz der musikalischen und kognitiven Fähigkeiten. Der fünfte Analyseschritt identifiziert über eine latente Klassenanalyse Gruppen von SuS, deren Wachstumskurven substantiell unterschiedlich verlaufen. Das Modell mit der besten Passung für die allgemeine musikalische Fähigkeit identifiziert drei latente Gruppen, die sich im allgemeinen Fähigkeitsniveau und ihren Wachstumsraten unterschieden. Bei allen drei Klassen wirkt CCM signifikant positiv auf die musikalische Entwicklung, wobei die Klassen mit dem niedrigsten und höchsten Niveau mehr von CCM profitieren. Die musikalischen Entwicklungsverläufe sind außerdem signifikant mit höherer Intelligenz, größerer Arbeitsgedächtniskapazität und höherem musikalischen Ausbildungsstand bei Studieneintritt assoziiert.

F-18: Motivation und Emotion von Studierenden im Hochschulkontext – Entwicklungstendenzen, Auswirkungen und Förderstrategien einer positiven akademischen Entwicklung

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* T03 R02 D81

Motivation und Emotion von Studierenden im Hochschulkontext – Entwicklungstendenzen, Auswirkungen und Förderstrategien einer positiven akademischen Entwicklung

Chair(s): **Luise von Keyserlingk** (TU Dortmund, Deutschland), **Charlott Rubach** (Universität Rostock, Deutschland)
*Diskutant*in(nen):* **Oliver Dickhäuser** (Universität Mannheim, Deutschland)

Studierende, die motiviert im Studium sind und sich wohlfühlen, zeigen häufig höhere Studienerfolge (Bäulke et al., 2018). Problematisch ist daher, dass empirische Befunde darauf hindeuten, dass die Motivation von Studierenden in der Studieneingangsphase stark absinkt (Benden & Lauermann, 2022) und immer mehr Studierende von psychischen Belastungen betroffen sind (Deutscher Bundestag, 2021; von Keyserlingk et al., 2020). Um diesem Problem kompetent zu begegnen und auf Seiten der Hochschule zu reagieren, sind empirische Befunde zu Bedingungsfaktoren dieser negativen Entwicklungstendenzen sowie effektiver intervenierender Strategien wünschenswert. Zudem sind weitere Forschungsbefunde erforderlich, die motivationale und emotionale Entwicklungstendenzen und Auswirkungen auf Leistungsvariablen replizieren - so können standortübergreifend Hochschulstrategien zur Unterstützung von Studierenden entwickelt und implementiert werden. Das vorgestellte Symposium führt vier Beiträge zusammen, die das Ziel haben, zwei Forschungsfragen zu untersuchen:

(1) Welche motivationalen und emotionalen Entwicklungstendenzen zeigen sich bei Studierenden und welche leistungsbezogenen Auswirkungen werden deutlich?

(2) Welche Bedingungsfaktoren zeigen bedeutsame Auswirkungen auf motivationale und emotionale Entwicklungstendenzen von Studierenden?

Zwei Beiträge befassen sich mit der ersten Forschungsfrage. In Beitrag 1 wurden Entwicklungsverläufe von Burnout und Zusammenhänge dieser mit motivationalen Überzeugungen bei Studierenden untersucht. Hier konnte ein linearer Anstieg von Burnout-Symptomen nachgewiesen werden, der in Abhängigkeit motivationaler Überzeugungen interindividuell unterschiedliche Tendenzen aufwies. Beitrag 2 nutzte einen innovativen Ansatz mit dem Ziel, Lernverhalten von Studierenden in Online-Kursen durch Klickverhalten im verwendeten Lernmanagement zu beschreiben und Zusammenhänge mit motivationalen Überzeugungen, Leistung und Stresserleben zu untersuchen. Studierende mit regelmäßigem Lernverhalten über den Semesterverlauf erzielten bessere Leistungen. Motivationale Überzeugungen waren zum Teil mit regelmäßigem Lernverhalten assoziiert und gingen zudem mit besseren Kursleistungen und geringerem Stresserleben einher.

Zwei Beiträge befassen sich mit der zweiten Forschungsfrage des Symposiums. Beitrag 3 untersuchte motivationale Entwicklungstrends von Studierenden mit dem Ziel, reziproke Effekte mit der Befriedigung psychologischer Grundbedürfnisse und selbstreguliertem Lernverhalten zu untersuchen. Befunde verwiesen darauf, dass Studierende mit besserem selbstregulierten Lernverhalten höhere intrinsische Motivation über die Zeit berichteten. Aufgrund der aufgezeigten negativen motivationalen und emotionalen Entwicklungstendenz von Studierenden untersuchte Beitrag 4 die Effektivität eines Interventionsprogramms mit Kontroll-Gruppen-Design mit dem Ziel,

negativen motivationalen und emotionalen Entwicklungstendenzen entgegenzuwirken. Das entwickelte Interventionsprogramm reduzierte negative Entwicklungstendenzen von Wertüberzeugungen im Studium und führte zur vermehrten Nutzung von Motivationsregulationsstrategien.

Die vier Beiträge leisten einen wichtigen Beitrag zum Verständnis der emotionalen und motivationalen Entwicklung von Studierenden an Hochschulen und tragen erste Befunde zusammen, die Hinweise auf Fördermöglichkeiten für eine positive akademische Entwicklung dieser geben. Das Symposium wird von Prof. Oliver Dickhäuser diskutiert.

Beiträge des Symposiums

Längsschnittliche Entwicklung von Burnout-Symptomen bei Studierenden und der Zusammenhang mit Studienmotivation aus Perspektive der Erwartung-Wert-Kosten Theorie

Derya Turhan[1], Anne Scheunemann[2], Theresa Schnettler[1], Lisa Bäulke[3], Daniel O. Thies[2], Markus Dresel[3], Stefan Fries[4], Detlev Leutner[5], Joachim Wirth[2], Carola Grunschel[1]
[1]Westfälische Wilhelms-Universität Münster, [2]Ruhr-Universität Bochum, [3]Universität Augsburg, [4]Universität Bielefeld, [5]Universität Duisburg-Essen

Theoretischer Hintergrund:

In Anbetracht vielfältiger Herausforderungen im Studium besteht für Studierende ein erhöhtes Risiko für das Erleben von Burnout-Symptomen wie emotionale Erschöpfung, Zynismus und ein reduziertes fachliches Wirksamkeitserleben (Schaufeli et al., 2002). Etablierte theoretische Modelle zu akademischem Burnout nehmen an, dass sich diese Symptome schrittweise über die Zeit entwickeln (Demerouti et al., 2001). Empirische Befunde zum zeitlichen Verlauf von Burnout-Symptomen bei Studierenden sind bisher begrenzt und inkonsistent. Einige Studien stellten einen linearen Anstieg und andere Studien hingegen keine wesentlichen Veränderungen der Burnout-Symptome über Jahre fest (Rudman & Gustavsson, 2012). Generell können solche Untersuchungen in Jahresintervallen nuancierte Veränderungen in Burnout-Symptomen übersehen, die potenziell innerhalb eines Semesters auftreten können.

Ein sukzessiver Verlust von Ressourcen spielt theoretischen Modellen zufolge eine zentrale Rolle für die Entwicklung dieser Symptome (vgl. Demerouti et al., 2001; Hobfoll, 1989). Da die verfügbaren Ressourcen von Studierenden typischerweise variieren, sind inter-individuelle Unterschiede in der Ausprägung von Burnout-Symptomen denkbar. Querschnittsstudien liefern erste Hinweise, dass diese Unterschiede unter anderem auf die Studienmotivation wie etwa auf den intrinsischen Wert von Aufgaben zurückzuführen sind (Chang et al., 2015; Lyndon et al., 2017). Eine differenzierte Betrachtung der Studienmotivation erfolgte in diese Studien nicht (z.B. Betrachtung der Erfolgserwartung und motivationalen Kosten). Darüber hinaus lassen die querschnittlichen Befunde keine Rückschlüsse auf die prädiktive Rolle der Studienmotivation für die Entwicklung von Burnout-Symptomen zu. Diese differenzierten Erkenntnisse sind insbesondere für die Gestaltung zugeschnittener Interventionsmaßnahmen erforderlich.

Fragestellungen und Methode:

Ziel der vorliegenden Längsschnittstudie mit drei Messzeitpunkten (Beginn, Mitte und Ende der Vorlesungszeit) war es, die aufgezeigten Forschungslücken zu adressieren. Zunächst modellierten

wir unkonditionale latente Wachstumskurven zweiter Ordnung, um die Ausgangsniveaus und einen möglichen linearen Anstieg der Burnout-Faktoren (emotionale Erschöpfung, Zynismus und reduziertes Wirksamkeitsvermögen) zu untersuchen. Anschließend überprüften wir mittels konditionaler latenter Wachstumskurven zweiter Ordnung, inwiefern diese Ausgangniveaus und Anstiege durch die Studienmotivation aus der Perspektive der Erwartung-Wert-Kosten Theorie (Eccles et al., 1983) vorhergesagt werden kann.

Zur Überprüfung der Fragestellungen wurden Längsschnittsdaten von N = 1435 Studierenden im Wintersemester 2018/2019 analysiert. Studierende waren in verschiedene Studienfächer immatrikuliert (MINT, Geistes-, Rechts- und Wirtschaftswissenschaften). Sie beantworteten Aussagen in Bezug auf die Symptome eines Burnouts in Form von emotionaler Erschöpfung (McDonald's ω = .77–.79), Zynismus (McDonald's ω = .82–.87) und reduziertem fachlichen Wirksamkeitserleben (McDonald's ω = .71–.74) sowie zu Studienmotivation in Form von Erfolgserwartung (McDonald's ω = .80), intrinsischem Wert (McDonald's ω = .90), persönlicher Wichtigkeit (McDonald's ω = .89), Opportunitätskosten (McDonald's ω = .84), Anstrengungskosten (McDonald's ω = .81) und emotionalen Kosten (McDonald's ω = .87).

Ergebnisse:

Es konnte ein linearer Anstieg aller Burnout-Symptomen (emotionale Erschöpfung, Zynismus und reduziertes fachliches Wirksamkeitserleben) im Semesterverlauf ermittelt werden. Zudem zeigten sich signifikante inter-individuelle Unterschiede in den Ausgangwerten und in den linearen Anstiegen der Burnout-Symptome. Zum Teil waren diese Unterschiede auf Facetten der Studienmotivation zurückzuführen. Insbesondere die Erfolgserwartung, der intrinsische Wert, die persönliche Wichtigkeit und die emotionalen Kosten erwiesen sich als bedeutsame Prädiktoren für den Ausgangswert von Burnout-Symptomen zu Semesterbeginn und den zeitlichen Verlauf von Burnout-Symptomen innerhalb eines Semesters.

Diskussion:

Die Ergebnisse implizieren, dass die längsschnittliche Erfassung von Burnout-Symptomen essentiell ist und das mehrere Facetten der Studienmotivation eine signifikante prädiktive Rolle für die Entwicklung von Burnout-Symptomen spielen und vor diesem Hintergrund im Rahmen der Prävention und Beratung dieser Symptome berücksichtigt werden sollten.

Lernverhalten und Motivation als Prädiktoren von Leistung und mentaler Gesundheit – Einblicke in den Hochschulkontext anhand von digitalen Verhaltensspurdaten und Surveydaten

Luise von Keyserlingk[1], Fani Lauermann[1], Renzhe Yu[2], Qiujie Li[2], Charlott Rubach[3], Richard Arum[2], Jacquelynne S. Eccles[2]
[1]TU Dortmund University, [2]University of California, Irvine, [3]Universität Rostock

Theoretischer Hintergrund

Das Studium ist eine essenzielle Qualifikationsphase zum Aufbau akademischer Fähigkeiten und domänenspezifischen Wissens für den späteren Beruf. Zugleich wird der Übergang an die Universität von Studierenden oft als herausfordernd erlebt und aktuelle Statistiken belegen eine erhöhte Stressbelastung und Beeinträchtigung der mentalen Gesundheit vieler Studierender (Stolzenberg et

al., 2020). Bisherige Literatur zeigt, dass motivationale Überzeugungen, wie Fähigkeitsüberzeugungen und subjektive Wertüberzeugungen in Kursen relevante Prädiktoren von Lernerfolg darstellen (Situated-Expectancy-Value-Theory; Eccles & Wigfield, 2020). Zudem ist die Fähigkeit selbstreguliert zu lernen eine wichtige Ressource für erfolgreiches Lernen an der Universität (Zimmerman & Schunk, 2011). Allerdings bringen Studierende zu Beginn des Studiums häufig erst wenig Erfahrungen im selbstregulierten Lernen mit (Cogliano et al., 2022). In dieser Studie wurde untersucht, welche Rolle motivationale Überzeugungen und Lernverhalten sowohl für den Leistungserfolg, als auch die mentale Gesundheit von Studierenden spielen.

Fragestellung

In den letzten Jahren wurden in der Hochschulforschung neben Selbstberichtdaten zu Lernstrategien auch vermehrt digitale Verhaltensspurdaten von Lernmanagement Systemen verwendet um Lernstrategien, Motivation und Lernerfolge von Studierenden zu untersuchen (Baker et al., 2020). In der vorliegenden Studie wurde anhand verschiedener Datenquellen untersucht, inwiefern Lernverhalten und motivationale Überzeugungen akademische und psycho-emotionale Outcomes von Studierenden vorhersagen.

Die spezifischen Fragestellungen lauteten:

a) Sind Verläufe von wöchentlichen Lernaktivitäten (hier: Klickaktivitäten in digitalen Verhaltensspurdaten) prädiktiv für finale Kursnoten?

b) Sind kursspezifische Fähigkeitsüberzeugungen und subjektive Wertüberzeugungen prädiktiv für wöchentliche Lernaktivitäten, finale Kursnoten, Zufriedenheit und Beeinträchtigung mentaler Gesundheit?

Methode

In der Studie wurden digitalte Verhaltentsspurdaten sowie Selbstberichtdaten zu kursspezifischen motivationalen Überzeugungen von Studierenden in Chemievorlesungen im Wintersemester 2020 an einer US-amerikanischen Universität verwendet. Dabei lagen digitale Verhaltensspurdaten von 1.597 Studierenden (Stichprobe A) aus vier Chemievorlesungen im Wintersemester 2020 vor. In diesem Semester fanden die Vorlesungen auf Grund der Corona-Pandemie vollständig digital statt. Eine Substichprobe von 54 Studierenden (Stichprobe B) beantworteten in der Mitte und nach Abschluss der Lehrveranstaltung Fragebögen zu motivationalen Überzeugungen, Zufriedenheit und mentaler Gesundheit. Zur Beantwortung der Fragestellungen wurden Mehrebenenanalysen genutzt (Level 1: wöchentlichen Klickaktivitäten, Level 2: Studierende).

Ergebnisse und ihre Bedeutung

Ergebnisse zu Fragestellung A (Stichprobe A) zeigten, dass die wöchentliche Lernaktivität gemessen über Klickaktivitäten in den online Vorlesungskursen über die Wochen des Semesters hinweg abnahmen. Dabei stellten Wochen mit Zwischen- und Abschlussprüfungen Ausnahmen mit stark erhöhter Klickaktivität dar. In den Wochen nach den Zwischenprüfungen war die Klickaktivität von Studierenden deutlich reduziert. Studierende mit einem geringeren Abfall in den Klickaktivitäten und regelmäßigeren Klickaktivitäten erzielten besser Kursnoten am Semesterende.

Ergebnisse zu Fragestellung B (Stichprobe B) zeigten, dass Studierende mit höheren Fähigkeitsüberzeugungen in ihren Kursen einen Stärkeren Abfall in Klickaktivitäten nach den Zwischenprüfungen aufzeigten (Tabelle 3). Höhere Fähigkeitsüberzeugungen waren jedoch mit höheren Kursnoten, höherer Zufriedenheit und geringerer mentaler Belastung am Ende des Semesters assoziiert. Höhere positive subjektive Wertüberzeugungen (wahrgenommene Wichtigkeit, Nützlichkeit, Interesse) in den Kursen führten zu höherer Zufriedenheit und geringerer mentaler Belastung am Semesterende. Studierende mit höheren wahrgenommenen Kosten im Kurs zeigten regelmäßigere Klickaktivitäten über das Semester hinweg und berichteten eine höhere mentale Belastung am Semesterende.

Zusammenfassend ist festzuhalten, das objektive Maße zu Lernverhalten durch digitale Verhaltensspuren (hier wöchentliche Klickaktivitäten in online Kursen) vielversprechende Datenquellen liefern um erfolgreiches Lernverhalten abzubilden. Regelmäßigere Klickaktivitäten waren in diesem Fall prädiktiver für finale Kursnoten als kurzfristige Lernanstiege in Examenswochen. Zudem zeigten Ergebnisse mit Daten der Substichprobe B, dass kurspezifische motivationale Überzeugungen mit Verläufen in wöchentlichen Lernaktivitäten im Zusammenhang standen. Ein zentraler Befund dieser Studie ist weiterhin, dass kurspezifische motivationale Überzeugungen nicht nur für Lernverhalten und Kursnoten von Bedeutung waren, sondern auch die Zufriedenheit und mentale Belastung von Studierenden am Semesterende vorhersagten. Auf Grund der kleinen Stichprobengröße, müssen diese Befunde in zukünftigen Studien jedoch erweitert werden.

Längsschnittliche Effekte psychologischer Grundbedürfnisse auf Motivation und selbstreguliertes Lernen während COVID-19

Elisabeth R. Pelikan, Luisa Grützmacher, Katharina Hager, Julia Holzer, Selma Korlat, Martin Mayerhofer, Barbara Schober, Christiane Spiel, Marko Lüftenegger
Universität Wien

Theoretischer Hintergrund

Die Selbstbestimmungstheorie (Selfdetermination Theory; SDT) von R. Ryan und E. Deci (2020) postuliert, dass die psychologischen Grundbedürfnisse nach wahrgenommener Kompetenz, Autonomie und sozialer Eingebundenheit von entscheidender Bedeutung für intrinsische Motivation (IM) und positives Lernverhalten ist. Laut der SDT wird die Befriedigung dieser psychologischen Grundbedürfnisse durch das soziale Umfeld beeinflusst (Ryan & Deci, 2020). Dementsprechend ist anzunehmen, dass die Umstellung vom Präsenzunterricht auf das Lernen auf Distanz, welches im März 2020 im Zuge der Eindämmungsmaßnahmen von COVID-19 weltweit beinahe flächendeckend der Fall war, einen Einfluss auf die Befriedigung der psychologischen Grundbedürfnisse und infolgedessen auf die IM von Studierenden hatte (Bartholomew et al., 2011). Zusätzlich bedeutet ein Lernen auf Distanz den Wegfall oder zumindest die Einschränkung etablierter Strukturen, sodass Studierende in zunehmendem Maß ihren Alltag und ihr Lernen selbst organisieren und regulieren mussten (Klingsieck et al., 2012). Fähigkeiten zum selbstregulierten Lernen (SRL) gewannen daher an Bedeutung.

Querschnittliche Zusammenhänge zwischen den psychologischen Grundbedürfnissen und IM wurden in verschiedenen Studien und Kontexten etabliert (Bureau et al., 2021; Ntoumanis et al., 2021; Reeve, 2016; Reeve & Jang, 2006; Rodrigues et al., 2021; Van den Broeck et al., 2016). Die Befundlage

bezüglich IM und SRL ist jedoch weniger eindeutig (Mukhtar et al., 2018; Ning & Downing, 2010). Studien, die wechselseitige Wirkzusammenhänge zwischen allen drei Konstrukten über die Zeit untersuchen, liegen nicht vor.

Ziele und Fragestellung

Ziel der vorliegenden Studie was es daher, die längsschnittlichen und wechselseitigen Effekte zwischen der Befriedigung der Grundbedürfnisse von Studierenden, IM und SRL während des Lernens auf Distanz im Zuge von COVID-19 zu untersuchen. Ausgehend von bisherigen Befunden nahmen wir an, dass die Befriedigung der psychologischen Grundbedürfnisse zu jedem Messzeitpunkt mit IM positiv zusammenhängt. Darüber hinaus gingen wir davon aus, dass die Befriedigung der drei Grundbedürfnisse positiv auf die IM zum nächsten Messzeitpunkt wirkt. Zusätzlich untersuchten wir exploratorisch wechselseitige Zusammenhänge zwischen IM und SRL, da in bisherigen Studien keine klare Wirkrichtung zwischen diesen beiden Konstrukten etabliert werden konnte.

Methode

Die Datenerhebung fand im Rahmen einer größeren Studie statt, in der österreichische Studierende über ihr Lernen und Wohlbefinden während COVID-19 befragt wurden. Insgesamt beantworteten N = 3286 Studierende (72,0 % weiblich, 27,2 % männlich, 0,4 % divers, 0,5 % fehlend) zwischen 14 – 71 Jahren (Mage = 25,06; SDage = 7,038) zwischen April 2020 und Juli 2021vier Online-Fragebögen. Für alle verwendeten Skalen (wahrgenommene Kompetenz, wahrgenommene Autonomie, wahrgenommene soziale Eingebundenheit, IM und SRL) konnte partielle metrische Messinvarianz über alle vier Messzeitpunkte nachgewiesen werden. Für die Überprüfung der beschriebenen Hypothesen wurde ein Cross-Lagged-Panel-Modell berechnet.

Ergebnisse und Bedeutung

Hypothesenkonform korrelierten die drei psychologischen Grundbedürfnisse zu jedem Messzeitpunkt positiv mit IM (r = .232 bis r = .641). Wider Erwarten wurden jedoch keine signifikanten Effekte auf IM über die Zeit gefunden. Bezüglich SRL zeigten sich über alle Messzeitpunkte hinweg durchgängig kleine, aber signifikanten positive Effekt auf IM. Im Gegensatz dazu wurde zwischen dem dritten und vierten Messzeitpunkt ein kleiner und – überraschender Weise – negativer signifikanter negativer Effekt von IM auf SRL gefunden.

Bedeutung

Obwohl die bereits etablierten querschnittlichen Assoziationen zwischen Autonomie, Kompetenz und sozialer Eingebundenheit bestätigt wurden, zeigt unsere Studie klar den Bedarf an längsschnittlichen Untersuchungen der in der SDT postulierten Zusammenhänge auf. Darüber hinaus wurde die Bedeutung von SRL für IM im Kontext des Lernens auf Distanz verdeutlicht. Weiters zeigte sich kein beziehungsweise sogar ein kleiner negativer Effekt von IM auf SRL. Eine Erklärung dafür könnte sein, dass Studierende mit hoher intrinsisch Motivation ihre Ziele ohne den bewussten Einsatz von SRL-Strategien verfolgen. Ein personenzentrierter Ansatz könnte weitere Hinweise auf die (Co-)Entwicklung und gegenseitige Beeinflussung von IM und SRL geben.

(Help me to) motivate me: Entwicklung und erste Bewährung eines Trainingsprogramms zur kombinierten Förderung von Motivation und Motivationsregulation im Studium

Sophie von der Mülbe, Markus Dresel
Universität Augsburg

Absinkende Studienmotivation ist ein bei Studierenden weit verbreitetes Phänomen gerade im ersten Studienjahr, das schwerwiegende Konsequenzen für Lernprozesse und deren Selbstregulation, Studienleistung, Wohlbefinden und Studienzufriedenheit haben und bis zum Studienabbruch führen kann (Dresel & Grassinger, 2013; Bäulke et al., 2018; Bäulke et al., 2021). Sie kann häufig auf eine mangelnde Selbstregulation der Motivation im Studium aufgrund fehlender Kompetenzen zur Motivationsregulation zurückgeführt werden (Steuer et al., 2019). Insbesondere im MINT-Bereich indizieren hohe Studienabbruchsquoten einen Bedarf für Trainingsprogramme zur Förderung von Motivation und Motivationsregulation (Heublein & Schmelzer, 2018; Heublein et al., 2017). Bisherige Trainingsansätze zielen rein auf die Förderung von Motivation (z.B. Perry et al., 1993; Hulleman et al., 2010) oder fokussieren auf Motivationsregulation (Eckerlein, 2020). Besonders sinnvoll erscheint jedoch die Kombination einer unmittelbaren Förderung von Motivation mit der Förderung von Motivationsregulation, um direkt gegen sinkende Motivation zu intervenieren und gleichzeitig Studierende zu befähigen, selbstreguliert mit auftretenden Motivationsproblemen umzugehen. Daher war das Ziel der vorliegenden Studie, ein Trainingsprogramm zur Förderung von Erfolgserwartung und Wert im Studium sowie zur Vermittlung motivationsregulatorischer Strategien zu entwickeln und im MINT-Bereich durchzuführen. Neben proximalen Effekten des Trainings auf Erwartung, Wert und den Einsatz von Motivationsregulationsstrategien im Studium wurden distale Effekte auf Prokrastination, Studienzufriedenheit und Studienabbruchsintentionen angestrebt.

Um das Training auf seine diesbezügliche Wirksamkeit zu prüfen, wurden alle Studierenden, die eine Informatik-Vorlesung für das zweite Fachsemester besuchten, zu Semesterbeginn und nach Semesterende zu einer Befragung zu ihrem Studierverhalten und -erleben eingeladen (Prä- und Posttest; ein dritter Messzeitpunkt, der als Follow-up fungiert, wird im Oktober 2022 realisiert - seine Ergebnisse werden bis zur Konferenz in den Beitrag integriert). Insgesamt liegen von 62 Studierenden die Daten aus Prä- und Posttest vor. 37 dieser Studierenden waren zufällig der Trainingsgruppe zugeteilt worden – sie nahmen über den Verlauf des Semesters an neun digitalen, ca. 30-minütigen Trainingseinheiten teil, in denen sie sich mit ihren Erfolgserwartungen und dem Wert ihres Studiums auseinandersetzten und den situationsgerechten Einsatz von Motivationsregulationsstrategien im Studium einübten. Die übrigen 25 Studierenden wurden der Kontrollgruppe zugeordnet.

In den vorläufigen Analysen mit den Daten aus Prä- und Posttest zeigten sich teilweise die erwarteten Effekte des Trainings. Für den subjektiven Wert des Studiums erwies sich die Interaktion zwischen Zeit- und Gruppenfaktor als statistisch signifikant ($F(1, 60) = 6.56$, $p < .01$, partielles $\eta^2 = 0.10$): Während der subjektive Wert des Studiums in der Kontrollgruppe vom ersten zum zweiten Messzeitpunkt deutlich absank, stieg er in der Trainingsgruppe an. Beim Einsatz von Motivationsregulationsstrategien ergab sich für die Strategie der Steigerung der persönlichen Bedeutsamkeit – neben Teilziele setzen, Selbstbelohnung und Fähigkeitsbezogener Selbstinstruktion eine der vier adressierten Strategien – eine statistisch signifikante Interaktion zwischen Zeit- und

Gruppenfaktor ($F(1, 60) = 8.16$, $p < .01$, partielles $\eta^2 = 0.12$): Der Einsatz dieser Strategie sank in der Kontrollgruppe ab, in der Trainingsgruppe stieg er über das Semester an. Signifikante Interaktionseffekte ergaben sich auch im Bereich der Studienzufriedenheit für die Zufriedenheit mit den Studieninhalten ($F(1, 60) = 4.96$, $p < .05$, partielles $\eta^2 = 0.08$) und der Zufriedenheit mit der Bewältigung von Studienbelastungen ($F(1, 60) = 5.32$, $p < .05$, partielles $\eta^2 = 0.08$). Während die Zufriedenheit mit den Studieninhalten und mit der Bewältigung von Studienbelastungen sich in der Kontrollgruppe über das Semester verschlechterten, entwickelten sich diese Komponenten der Studienzufriedenheit in der Trainingsgruppe positiv. Für die Erfolgserwartung, Prokrastination und die Studienabbruchsintentionen waren keine Trainingseffekte statistisch abzusichern.

Insgesamt liefern die vorläufigen Ergebnisse Hinweise darauf, dass das Training zumindest teilweise die angestrebten Wirkungen auf das Erleben im Studium hat und auf motivationaler Ebene besonders die Wert-Komponente der Motivation anzusprechen scheint. Dies indiziert, dass ein Training, dass eine unmittelbare Förderung von Motivation mit einer Förderung der Selbstregulation der Motivation kombiniert, ein grundsätzlich lohnender Ansatz sein könnte.

F-19: Survey-Methoden in der Hochschulforschung

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* R11 T00 D03

Survey-Methoden in der Hochschulforschung

Chair(s): **Gesche Brandt** (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Germany), **Susanne de Vogel** (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Germany)
*Diskutant*in(nen):* **Gesche Brandt** (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Germany), **Susanne de Vogel** (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Germany)

Die von der Hochschulforschung untersuchte Gruppe der Hochqualifizierten ist in vielerlei Hinsicht besonders. Zum einen hebt sie sich hinsichtlich ihrer demographischen Zusammensetzung und ihrer Lebenssituation deutlich vom Durchschnitt der Gesamtbevölkerung ab (hochgebildet, berufsorientiert, räumlich mobil). Zum anderen ist sie in sich sehr heterogen, denn sie umfasst Personen, die sich in unterschiedlichen Qualifikationsstufen, fachlichen Kontexten und Beschäftigungssituationen befinden. Das wissenschaftliche Interesse an der Erforschung von verschiedenen Personengruppen des tertiären Bildungsbereichs ist in den vergangenen Jahren – u. a. aufgrund der Expansion des tertiären Sektors und des zunehmenden Bedarfs an Hochqualifizierten – stark gestiegen.

Zur Erfassung langjähriger Bildungs- und Berufsverlaufsdaten und der Erforschung kausaler Zusammenhänge sind insbesondere Panelbefragungen in der Hochschulforschung unverzichtbar. Die zunehmende Internationalisierung des Wissenschaftssystems und die von hochschulpolitischen Akteuren forcierte Schaffung eines Europäischen Hochschulraums erzeugen zusätzlich einen Bedarf an internationalen Vergleichsstudien. Dabei müssen sich auch nationale Befragungen verstärkt auf eine internationale Zielgruppe einstellen.

Somit stehen Hochschulforschende bei der Konzeption und Durchführung von Primärerhebungen vor vielfältigen Herausforderungen. Bei der Erstellung des Befragungsprogramms von Primärerhebungen stehen sie vor der Frage, inwiefern Erhebungsinstrumente aus allgemeinen Bevölkerungsumfragen für die sehr spezifische Gruppe der Hochqualifizierten geeignet und auf die zugleich heterogene Gruppe unterschiedlicher Disziplinen anwendbar sind. Außerdem stellt sich die Frage, wie man die spezielle Zielgruppe der Hochqualifizierten am zuverlässigsten erreicht und mit welchen Mitteln man sie für eine (regelmäßige) Teilnahme an (Panel-)Befragungen motivieren kann.

Der erste Beitrag zeigt die Entwicklung eines spezifischen Instruments, das eigens für die Bedarfe innerhalb der Bildungsforschung entwickelt wurde. In der Schweizer Entwicklungsstudie wird die Konstruktion und Validierung eines Instruments zur Erfassung der Studierfähigkeit beschrieben, wobei zur Entwicklung des Instruments in neuartiger Weise die Lehrmaterialen der Hochschulen herangezogen und dabei auch fachspezifische Besonderheiten berücksichtigt werden.

Der zweite Beitrag befasst sich mit der Nutzungsmöglichkeit eines Instruments für die Hochschulforschung, das für allgemeine Bevölkerungsumfragen entwickelt wurde. Mit Daten der DZHW-Absolventenbefragung wird ein Kurzinstrument zur Messung generalisierter Umfrageeinstellungen validiert und dessen Übertragbarkeit auf die Gruppe der Hochschulabsolvent*innen diskutiert.

Im dritten Beitrag werden die Determinanten des Panel-Dropouts im Magdeburger Studierendenpanel analysiert und gezeigt, welche personen- und befragungsbezogenen Faktoren Unit-Nonresponse speziell in Studierendenbefragungen bedingen.

Der vierte Beitrag befasst sich mit der Konzeption und Umsetzung des German International Student Survey und liefert Erkenntnisse zu den Determinanten der Panelmortalität in international zusammengesetzten Befragungsgruppen.

Die Beiträge liefern wichtige methodische Impulse für die Hochschulforschung, aber auch für andere empirische Forschungsfelder. Sie zeigen auf, wie Erhebungen und Survey-Instrumente für spezielle Zielgruppen entwickelt und angepasst werden können.

Beiträge des Symposiums

Entwicklung eines passgenauen Instruments zur Erfassung allgemeiner Studierfähigkeit anhand einer neuartigen Methodik

Maren Oepke[1], **Karin Gehrer**[2]
[1]Universität Zürich, [2]Leibniz-Institut für Bildungsverläufe, Bamberg

Im Kontext einer zunehmenden Kompetenzorientierung in den letzten Jahren auch für den Bereich des Hochschulsektors (Wissenschaftsrat, 2015) ist die Diagnostik von fachlichen und fachübergreifenden Kenntnissen, Fähigkeiten und Fertigkeiten Studierender in den aktuellen Forschungsfokus gerückt (Zlatkin-Troitschanskaia et al., 2016). Dies gilt insbesondere vor dem Hintergrund der Heterogenität der Studierenden i.S. von großen Unterschieden der Studieneingangsvoraussetzungen, die dramatisch zugenommen haben (Köller, Pant & Zlatkin-Troitschanskaia, 2020).

Dabei wird Heterogenität vielfach mit mangelnder Studierfähigkeit gleichgesetzt und als Defizit der Studierenden angesehen, dem mit verschiedenen Maßnahmen wie Beratungsangeboten, Brückenkursen oder Tutorien im ersten Studienjahr begegnet werden soll (Hanft, 2015; Tieben, 2019). Um diese Maßnahmen passgenau anbieten zu können, wären objektive, valide und zuverlässige Methoden zur Diagnostik der Kompetenzen von Studierenden notwendig, um Unterschiede in den Eingangsvoraussetzungen frühzeitig aufdecken und angemessen berücksichtigen zu können (Wissenschaftsrat, 2015). Allerdings mangelt es im deutschsprachigen Raum an validierten Testinstrumenten für eine fachliche als auch fachübergreifende Eingangsdiagnostik von Studienanfänger*innen (Köller et al., 2020).

Hier setzt der vorliegende Beitrag an: Vorgestellt wird ein objektives und validiertes Testinstrument und seine Entwicklung zur Erfassung von Facetten allgemeiner Studierfähigkeit. Das Instrument wurde im Rahmen einer schweizweiten Studie zur Evaluation der letzten Maturitätsreform entwickelt, validiert und erfolgreich eingesetzt (EVAMAR II, Eberle et al., 2008). Studierfähigkeit wurde darin in Anlehnung an Huber (2009) definiert als ein Ensemble von Fähigkeiten, die es ermöglichen, ein Studium erfolgreich zu beginnen und durchzuführen. Die Entwicklung des Instruments, bestehend aus zwei Tests mit basalen erstsprachlichen und mathematischen Inhalten, greift die bisher kaum genutzte Methode auf, Studierfähigkeit aus Sicht der abnehmenden Institutionen zu erfassen, indem allgemeine Anforderungen von Hochschul-Lehrmaterialien detailliert analysiert wurden (Klomfaß, 2011; Huber, 2009).

Zur Operationalisierung des Konstruktes der „Studierfähigkeit“ und Konkretisierung der Anforderungen an die allgemeine Studierfähigkeit aus Sicht der Universitäten wurden neben einer Befragung von Hochschuldozierenden Inhaltsanalysen von Fachcurricula, Lehrbüchern, Skripten und

Vorlesungen durchgeführt. Dadurch konnten vorausgesetzte fachliche Anforderungen verschiedener Studienfächer an das Wissen der Studienanfänger*innen ermittelt und festgehalten werden. Dieses sogenannte „Eingangswissen“ musste von den erstsemestrigen Studierenden zuvor am Gymnasium erlernt worden sein.

Untersucht wurden die Studienmaterialien im Hinblick auf die Identifikation von Sinneinheiten (Begrifflichkeiten, Modelle oder Theorien), die für das Verständnis der Inhalte der Materialien notwendig waren und gleichzeitig nicht in den Materialien erklärt wurden. Die Analyse erfolgte mittels einer fein- und einer grobmaschigen Kodierung. Während sich die grobmaschige Kodierung überblickshaft auf größere Textteile bezog, handelte es sich bei der feinmaschigen Kodierung um eine kleinschrittige Methode mit vergleichsweise weniger umfassenderen Sinneinheiten. Die feinmaschige Kodierung des Lehrmaterials diente als Basis für die Konstruktion von Instrumenten zur Überprüfung des vorhandenen Eingangswissens bei Studienberechtigten.

Auf Basis der so entstandenen Datenbank des Eingangswissens aus den gymnasialen Schulfächern und dem Studienmaterial aus verschiedenen Studienfächern wurden Leistungstests für Erstsprache und Mathematik konstruiert. Dazu wurden Kompetenzraster entwickelt, die das Eingangswissen und die zu entwickelnden Items im Hinblick auf fachspezifische Fähigkeiten und Inhalte systematisierten. Die Leistungstests wurden in einer nahezu schweizweiten Erhebung beim Maturitätsjahrgang 2007 kurz vor der Matura eingesetzt und mithilfe der Rasch-Methode ausgewertet.

Positive und signifikante Zusammenhänge mit den im Rahmen der EVAMARII-Studie ebenfalls erhobenen Schulnoten sprechen für die externe Validität der beiden Leistungstests. Aufgrund Daten einer drei bis vier Jahre später stattfindenden Nachfolgestudie, in der für einen Teil der deutschsprachigen EVAMARII-Stichprobe Indikatoren für ihren Studienerfolg erhoben wurden (Oepke & Eberle, 2010), konnten die Werte der Leistungstests mit positivem Ergebnis auf ihre Prognosevalidität überprüft werden.

Neben der Darstellung der Methode und den Ergebnissen der Validierung wird im Beitrag abschließend diskutiert, inwiefern die beschriebenen Tests weiterhin eingesetzt werden können, ohne ihre Passgenauigkeit zu verlieren, und welche Implikationen sich für die Konstruktion und Nutzung eines solchen Verfahrens ergeben.

Validierung der Survey Attitude Scale zur Messung generalisierter Umfrageeinstellungen: Ergebnisse eines Survey-Experiments bei Hochschulabsolvent*innen

Thorsten Euler, Ulrike Schwabe, Isabell Fiedler
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, Germany

Der allgemeine Trend zur sinkenden Teilnahmebereitschaft an (wissenschaftlichen) Umfragen stellt auch für die quantitativ-empirische Hochschulforschung eine Herausforderung dar (Menold & Züll, 2011; Meyer et al., 2015; Ramm, 2014). Groß angelegte regelmäßige Befragungen von Hochqualifizierten bilden vor dem Hintergrund von Bildungsexpansion und der gestiegenen Bedeutung von Wissensarbeit auf dem Arbeitsmarkt eine wichtige Grundlage für Monitoring und Berichtspflichten einerseits als auch für die empirische Beantwortung aktueller Forschungsfragen andererseits. Auf Letzterem basieren idealerweise evidenzbasierte Entscheidungen in der Hochschul- und Arbeitsmarktpolitik.

Vor diesem Hintergrund erscheint eine Auseinandersetzung mit den Determinanten der Befragungsteilnahme notwendig (Konradt & Fary, 2006; Krieger, 2018; Meyer et al., 2015; Stoop et al., 2010). Deshalb haben de Leeuw et al. (2019) die internationale Survey Attitude Scale (SAS), ein Neun-Item-Instrument zur Erfassung generalisierter Umfrageeinstellungen, entwickelt. Diese Skala unterscheidet theoretisch und empirisch drei Einstellungsdimensionen: (i) Freude an Umfragen (Survey Enjoyment), (ii) Nutzen von Umfragen (Survey Value) und (iii) Belastung durch Umfragen (Survey Burden). Obwohl dieses Kurzinstrument sehr vielversprechend ist, wurde es ausschließlich mit Daten der allgemeinen Bevölkerung validiert. Zudem haben Analysen gezeigt, dass die beiden Dimensionen Survey Value und Survey Burden messtheoretische Schwächen wie Kreuzladungen und niedrige Werte für die interne Konsistenz aufweisen (de Leeuw et al., 2019; Fiedler et al., 2022).

Wir verwenden die von de Leeuw et al. (2019) entwickelte SAS, um erstens empirisch zu überprüfen, ob das internationale Kurzinstrument für die Gruppe der Hochqualifzierten einsetzbar ist. Zweitens untersuchen wir, inwiefern messtheoretische Schwächen des Originalinstruments durch Modifkationen in den konkreten Formulierungen der Items und/oder durch die Hinzunahme zusätzlicher, bereits in der empirischen Literatur diskutierter Items behoben werden können.

Zur Validierung der SAS haben wir ein Experiment in eine Befragung des DZHW Absolventenpanels 2009 integriert. Dessen Grundgesamtheit bilden alle Hochschulabsolvent*innen, die im Prüfungsjahr 2009 ihren ersten berufsqualifzierenden Studienabschluss an einer staatlich anerkannten Hochschule in Deutschland erworben haben (Baillet et al., 2017).

Für die Implementierung des Experimentaldesigns wurden alle Befragten ohne Kinder zu Beginn der Befragung randomisiert einer von zwei Gruppen zugewiesen („Split-Ballot-Design“: Groves et al., 2011).

In unseren Analysen werden diejenigen Befragten berücksichtigt, die auf allen Variablen gültige Angaben haben. Basierend auf der zufälligen Zuweisung wurden 664 Befragte der Kontrollgruppe und 624 Befragte der Experimentalgruppe zugewiesen. Zur Überprüfung der internen Validität des Kurzinstruments für generalisierte Umfrageeinstellungen rechnen wir konfirmatorische Faktorenanalysen (CFA) und berichten neben den (nicht standardisierten) Faktorladungen selbst Maße für die Modellgüte unserer Schätzungen.

Die empirischen Ergebnisse zeigen zunächst, dass sich die theoretisch angenommene dreidimensionale Struktur der SAS für die Gruppe der Hochqualifzierten empirisch abbilden lässt. Zudem finden wir in unseren Analysen, dass punktuelle sprachliche Modifkationen zur Verbesserung der Reliabilität und Modellpassung beitragen. Die Hinzunahme von drei zusätzlichen Items führt hingegen nicht wie angenommen zu einer Verbesserung des Gesamtmodells. Insofern hat sich die sparsame Messung der drei latenten Dimensionen mit je drei manifesten Items bewährt (de Leeuw et al., 2019). Dies wird in unseren Analysen durch die erneute Reduktion auf neun Items bestätigt.

In der experimentellen Weiterentwicklung einer solchen Skala sehen wir großes Potenzial für die Integration der Messung stabiler Dispositionen der Befragten in thematische Surveys im Bereich der empirischen Hochschulforschung. Zusammengenommen leisten wir damit einen Beitrag zur Validierung und Verbesserung der SAS. Nicht zuletzt zeigt diese Studie aber auch ganz konkret auf,

wie Methodenforschung und inhaltlich-orientierte empirische Hochschulforschung stärker miteinander verzahnt werden können.

Die Problematik von Unit-Nonresponse in Panelbefragungen: Eine Analyse anhand der Daten des Studierendenpanels der Otto-von-Guericke-Universität Magdeburg

Sarah Berndt, **Annika Felix**, **Anke Manukjan**
Otto-von-Guericke Universität Magdeburg

Die Anzahl der Erhebungen an deutschen Hochschulen nimmt immer mehr zu, wobei insbesondere die Studierenden als Adressat*innen von Befragungen gelten. Von besonderer Bedeutung für die Qualitätssicherung an Hochschulen sind dabei längsschnittliche studentische Sichtweisen. Durch sinkende Rücklaufquoten und eine sogenannte „Befragungsmüdigkeit“, die sich auch im universitären Kontext abzeichnet (vgl. Hauss & Seyfried, 2019, S. 10f.), besteht jedoch die Gefahr, dass ihr Nutzen unterminiert wird. Der gänzliche Ausfall von Teilnehmenden (Unit-Nonresponse) (vgl. Schnell et al., 2005, S. 306) wird zur Fehlerquelle von Befragungsergebnissen, wenn sich die Ausfälle nicht zufällig, sondern systematisch verteilen und folglich eine Verzerrung entsteht (Nonresponse-Bias) (vgl. Koch & Blohm, 2015, S. 1). Die Möglichkeit für Unit-Nonresponse besteht in Panelbefragungen in jeder Welle neu, sodass sich Verzerrungen der Daten folglich als Konsequenz von Initial Wave Nonresponse (d. h. Ausfällen in der ersten Befragungswelle) sowie von Panelmortalität (d. h. Ausfällen in Folgewellen) ergeben können (vgl. Hoch, 2013, S. 8; Schupp, 2014, S. 934f.).

Es existieren verschiedene theoretische Zugänge, um die Teilnahmebereitschaft der Zielpersonen an Befragungen zu erklären. Häufig werden handlungstheoretische Modelle wie die Wert-Erwartungstheorie nach Esser (vgl. Esser, 1986), die Theorie des sozialen Austauschs (vgl. Homans, 1958), die Theorie der sozialen Isolation (vgl. Groves & Couper, 1998) oder die Leverage-Salience-Theorie (vgl. Groves et al., 2000) herangezogen. Eine weitere theoretische Perspektive eröffnen sozialpsychologische Theorien, die auf Einstellungen und Verhalten rekurrieren, wie die Theorie des geplanten Verhaltens (vgl. Ajzen, 1985).

Der Beitrag geht in diesem Kontext anhand des Studierendenpanels der Otto-von-Guericke-Universität Magdeburg (OVGU) der Frage nach, ob und welche personen- und befragungsbezogenen Faktoren Unit-Nonresponse in Studierendenbefragungen bedingen, um aus den Befunden Handlungsmöglichkeiten abzuleiten. Zur Analyse der systematischen Ausfälle kommen zwei Verfahren zum Einsatz. Zum einen wird der Ausfall zu Beginn der Panelerhebung über den Vergleich mit Aggregatstatistiken betrachtet, zum anderen wird er zwischen den Erhebungswellen 1 und 2 sowie 1 und 3 untersucht.

Ein Abgleich der Daten des Studierendenpanels mit den Verteilungen der Hochschulstatistik zeigt bezüglich des Teilnahmeverhaltens in der jeweils ersten Welle der vier untersuchten Kohorten Verzerrungen in Hinblick auf die Fächergruppe. Es besteht zudem eine geschlechterspezifische Selektivität dergestalt, dass Frauen deutlich überrepräsentiert sind. Die bereits in der Erstsemesterbefragung bestehende Verzerrung verstärkt sich durch die höhere Teilnahmebereitschaft von Frauen an Folgewellen nochmals. Darüber hinaus zeigt die Betrachtung des systematischen Panelausfalls zwischen den Erhebungswellen mittels binär-logistischer Regressionsanalysen, dass weitere soziodemografische, studien-, persönlichkeits- und

befragungsbezogene Variablen einen Einfluss auf die Teilnahmewahrscheinlichkeit an Folgebefragungen haben. Dabei ist die Wahrscheinlichkeit zur Teilnahme mit steigendem Alter, bei in Deutschland geboren Personen (gegenüber Studierenden mit Migrationshintergrund) und bei leistungsstarken Studierenden sowie bei absteigender Ausprägung des Big Five Persönlichkeitsmerkmals Extraversion erhöht. Für die Teilnahme an der ersten Folgebefragung erweisen sich zusätzlich die Fächergruppe Wirtschaftswissenschaften (gegenüber den Humanwissenschaften) und die befragungsbezogenen Aspekte Vergabe von Incentives sowie die Anzahl der Erinnerungen als relevant. Die Teilnahme an Welle 3 wird hingegen ergänzend durch das Kommunikationskonzept und das Big Five Persönlichkeitsmerkmal Gewissenhaftigkeit bedingt. Dabei erhöht ein entsprechendes Kommunikationskonzept die Teilnahmewahrscheinlichkeit. Gleiches zeichnet sich für gewissenhafte Studierende ab. Die Ergebnisse decken sich weitgehend mit den Befunden aus allgemeinen Bevölkerungsumfragen. Differenzen bestehen hingegen gegenüber Untersuchungen zum Teilnahmeverhalten im Hochschulkontext.

Zusammenfassend zeigen die Auswertungen, dass vor der Analyse interessierender Variablen stets geprüft werden sollte, ob ein Zusammenhang zwischen diesen und den identifizierten Einflussgrößen besteht. Ist dies der Fall, ist wiederum ein Nonresponse-Bias in Bezug auf die interessierende Variable zu erwarten, welcher durch Gewichtungsverfahren korrigiert werden kann bzw. in jedem Fall in der Interpretation der Ergebnisse Berücksichtigung finden muss (vgl. Volk & Hochgürtel, 2016, S. 68f.). Um Verzerrungen aufgrund von systematischem Nonresponse in Onlinebefragungen zu vermeiden, sollten Präventivmaßnahmen ergriffen werden, welche die Ausfälle minimieren. Der Beitrag diskutiert entsprechende Möglichkeiten vor dem Hintergrund der empirischen Ergebnisse.

Einflussfaktoren des Antwortverhaltens von internationalen Studierenden in Studierendenbefragungen

Susanne Falk, **Theresa Thies**
Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung

Studierendenbefragungen sind für die Analyse der Qualität von Studiengängen, des Studienerfolgs und des Studienabbruchs für Hochschulen und Hochschulforschung von entscheidender Bedeutung. Die Zahl der internationalen Studierenden an deutschen Hochschulen steigt seit Jahren (DAAD und DZHW 2021) – trotzdem sind diese in repräsentativen Studierendenbefragungen unterrepräsentiert. Grund dafür ist, dass deren Rekrutierung aufgrund von Sprachbarrieren schwieriger ist als die Rekrutierung inländischer Studierender. In dem Beitrag wird untersucht, welche Faktoren eine Teilnahme an Studierendenbefragungen für internationale Studierende begünstigen oder erschweren. Es wird in Übereinstimmung mit früheren Studien (Zinn et al. 2020; Zinn et al. 2017; Liebeskind und Vietgen 2017) davon ausgegangen, dass besonders leistungsstarke Studierende eine höhere, bzw. zum Abbruch neigende Studierende eine niedrigere Wahrscheinlichkeit der Befragungsteilnahme haben. Insbesondere Studierende mit finanziellen Schwierigkeiten sollten eine höhere Teilnahmewahrscheinlichkeit aufweisen, wenn großzügige Verlosungen oder direkte monetäre Anreize geboten werden.

Beantwortet wird die Forschungsfrage mittels einer logistischen Regressionsanalyse zur Teilnahmewahrscheinlichkeit in den jeweils ersten vier Befragungswellen des International Student

Survey. Der Survey ist eine Online-Panelerhebung unter internationalen Bachelor- und Masterstudierenden an deutschen Hochschulen (N=4751). Die Ergebnisse zeigen, dass die Determinanten des Rücklaufs je nach untersuchter Panelwelle unterschiedlich sind. Das heißt, es gibt keine Merkmalsgruppe, die zu jedem Zeitpunkt und kontinuierlich aus der Studierendenbefragung herausfällt. Etwa 60% der Studierenden füllten den Fragebogen auf Englisch aus. Bessere Sprachkenntnisse in Englisch gingen mit einer höheren Wahrscheinlichkeit einher, die Befragungen der Wellen 2 und 3 zu beantworten. Eine bessere Hochschulzugangsberechtigungsnote und finanzielle Schwierigkeiten korrelierten positiv, und eine Studienabbruch- oder Fachwechselabsicht negativ mit der Antwortwahrscheinlichkeit in einzelnen Wellen. Die Ergebnisse sollen Personen an Hochschulen und in der Hochschulforschung helfen, selbst Befragungen für internationale Studierende zu konzipieren und deren Antwortverhalten zu verstehen. Darüber hinaus sollen sie dazu anregen, internationale Studierende in bestehende Studierendenbefragungen stärker einzubinden, beispielsweise durch das zusätzliche Angebot von englischsprachigen Fragebögen.

F-20: Evidenz in der empirischen Bildungsforschung: Auch eine Frage der Methode?

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* R11 T00 D05

Evidenz in der empirischen Bildungsforschung: Auch eine Frage der Methode?

Chair(s): **Nicolas Hübner** (Universität Tübingen, Germany)
*Diskutant*in(nen):* **Benjamin Nagengast** (Universität Tübingen, Germany)

In der empirischen Bildungsforschung wird bereits lange diskutiert, was als Evidenz verstanden werden kann und inwieweit Bildungsforschung überhaupt in der Lage ist, handlungsleitende Evidenzen für rationale Entscheidungen bereitzustellen (z.B. Baumert, 2016). In den entsprechenden Diskussionen wurde verschiedentlich auf die besondere Bedeutung methodologischer Kriterien für den aufwendigen Prozess der Evidenzgewinnung und Evidenzkategorisierung hingewiesen (Bromme, Prenzel & Jäger, 2014). Klassische Kategoriensysteme zur Beurteilung der Qualität von Forschungsergebnissen sehen insbesondere in randomisierten Experimenten und darauf basierenden Reviews und Meta-Analysen einen vielversprechenden Weg zur Generierung besonders belastbarer Evidenzen. Gleichzeitig sind Experimente für viele Fragestellungen der empirischen Bildungsforschung aus verschiedenen Gründen oft nicht durchführbar, und so finden sich insbesondere in der jüngeren Vergangenheit vermehrt Studien, die sich mit Möglichkeiten und Grenzen der Schätzung kausaler Effekte in Beobachtungsstudien beschäftigen. Auch zentrale Reviews und Meta-Analysen der Bildungsforschung greifen inzwischen in einem substantiellen Ausmaß auf Ergebnisse aus entsprechenden Beobachtungsstudien zurück. Das vorliegende Symposium zielt darauf ab, in insgesamt vier Beiträgen unterschiedliche methodische Zugänge sowie deren Möglichkeiten und Grenzen zur Evidenzgewinnung konzeptuell näher darzustellen. Die Bedeutung der Zugänge für das Spannungsfeld zwischen Unsicherheit und Evidenz werden abschließend genauer eingeordnet und diskutiert.

Im ersten Beitrag erfolgt zunächst eine kurze Einführung zum Thema der kausalen Effekte in Beobachtungsstudien, bevor anschließend prominente längsschnittliche Strukturgleichungsmodelle (Cross-Lagged Panelmodelle) näher in Bezug auf ihr Potenzial zur Identifikation kausaler Effekte diskutiert werden (Oliver Lüdtke und Alexander Robitzsch).

Im zweiten Beitrag werden Ergebnisse zur Untersuchung des Zusammenhangs von Leistung und Selbstkonzept mit traditionellen längsschnittlichen Strukturgleichungsmodellen (mit und ohne Kovariaten) und Weighting-Verfahren (z.B. Entropy balancing) kontrastiert (Nicolas Hübner, Wolfgang Wagner, Steffen Zitzmann & Benjamin Nagengast).

Der dritte Beitrag beschäftigt sich schließlich mit den Effekten der G8-Reform auf das Lernverhalten von Studierenden und vergleicht hierzu Ergebnisse unter Verwendung einer traditionellen Variante der Difference-in-difference-Schätzung mit einer neuen Variante für gestaffelte Reformeinführungen (Jakob Schwerter, Nicolai Netz & Nicolas Hübner).

Der vierte Beitrag widmet sich der Frage, in welchem Ausmaß Meta-Analysen Informationen zur Beurteilung der Qualität der Evidenz berücksichtigen. Dazu wird die Transparenz und Reproduzierbarkeit von Meta-Analysen im Forschungsfeld der Selbstkonzepte von Schülerinnen und Schülern untersucht (Thorben Jansen, Jennifer Meyer und Jens Möller).

Die Bedeutung der Ergebnisse der einzelnen Studien werden vor dem Hintergrund des Tagungsthemas eingeordnet und diskutiert (Benjamin Nagengast).

Beiträge des Symposiums

Schätzung von kausalen Effekten mit Hilfe von Cross-Lagged Panelmodellen: Ein Vergleich verschiedener Ansätze

Oliver Lüdtke, Alexander Robitzsch
IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Zentrum für internationale Bildungsvergleichsstudien (ZIB)

Theoretischer Hintergrund.

In der empirischen Bildungsforschung werden häufig Cross-Lagged Paneldesigns umgesetzt, in denen (mindestens) zwei Variablen (Xt und Yt) über die Zeit erhoben werden (siehe z.B. den Zusammenhang zwischen schulischer Leistung und Motivation; Marsh et al., 2005). Im Zentrum steht meistens der zeitversetzte Effekt (Cross-Lagged Effekt) einer Variable (Xt-1) auf eine andere Variable (Yt). Zur Schätzung des Cross-Lagged Effekts werden neben dem traditionellen Cross-Lagged Panelmodell (CLPM) auch zunehmend Strukturgleichungsmodelle eingesetzt, in denen durch die Spezifikation zusätzlicher latenter Variablen für die Effekte (unbeobachteter) zeitstabiler Faktoren (Stable Traits) kontrolliert werden soll. Zu nennen wären hier vor allem das Random-Intercept CLPM (RI-CLPM; Hamaker et al., 2015), das GCLPM (General CLPM; Zyphur et al., 2020) und ein CLPM mit Fixed Effects (FE-CLPM; Allison et al., 2017; Dishop & DeShon, 2021). In der Auswertung von Daten aus Cross-Lagged Paneldesigns stellt sich somit die Herausforderung, welches dieser Modelle zur Schätzung des Cross-Lagged Effekts gewählt werden soll – insbesondere, wenn die verschiedenen Modelle zu unterschiedlichen Schätzungen des Cross-Lagged Effekts führen. Der vorliegende Beitrag vergleicht die verschiedenen Ansätze aus einer kausalen Perspektive (Lüdtke & Robitzsch, 2022).

Fragestellung.

Eine notwendige Bedingung für die kausale Interpretation des Cross-Lagged Effekts ist, dass alle relevanten Kovariaten berücksichtigt werden (Hernán & Robins, 2020). Während das traditionelle CLPM davon ausgeht, dass diese Kovariaten beobachtet und in das Modell aufgenommen wurden, besitzen die alternativen Modelle (RI-CLPM, GCLPM und FE-CLPM) den Vorteil, dass mit Hilfe zusätzlicher Variablen auch für die möglichen Effekte von unbeobachteten (zeitstabilen) Kovariaten kontrolliert werden kann. Allerdings beruht die Identifikation der Effekte der latenten Variablen auf starken Annahmen über den zeitlichen Prozess der beiden Variablen (Xt und Yt).

Methode.

Mit Hilfe von vier Simulationsstudien wird illustriert, wie sensitiv die Schätzung von Cross-Lagged Effekten gegenüber unterschiedlichen Modellierungsannahmen ist. Es wird von einem Design mit zwei Variablen ausgegangen, die über drei Messzeitpunkte erhoben wurden (X1, X2, X3 und Y1, Y2, Y3). Im Zentrum steht die Schätzung des Cross-Lagged Effekts von X2 auf Y3. In Studie 1 ist ein CLPM mit zusätzlichen lag-2 Effekten (z.B. Effekte von X1 auf Y3) das wahre datengenerierende Modell. Studie 2 geht davon, dass es eine unbeobachtete Variable gibt, die die beobachteten

Variablen (X1,…, Y3) direkt beeinflusst. Studie 3 nimmt an, dass das wahre Modell ein RI-CLPM ist. In Studie 4 werden die Daten mit einem FE-CLPM nach Allison et al. (2017) generiert.

Ergebnisse und ihre Bedeutung.

Die Hauptbefunde der Simulation lassen sich wie folgt zusammenfassen. Erstens wird bestätigt, dass das CLPM zu verzerrten Schätzungen des Cross-Lagged Effekts führt, wenn nicht alle Variablen beobachtet wurden (Studie 2, 3 und 4). Zweitens zeigen die alternativen Ansätze, die mit Hilfe zusätzlicher latenter Variablen für die Effekte zeitstabiler Faktoren korrigieren, stark verzerrte Schätzungen, wenn die Modellannahmen verletzt sind. Beispielsweise führt das RI-CLPM zu verzerrten Schätzungen des Cross-Lagged Effekts in den Studien 1, 2 und 4. Drittens zeigen die Befunde, dass das FE-CLPM mit weniger starken Modellierungsannahmen verbunden ist als das RI-CLPM oder das GCLPM. Dies liegt darin begründet, dass das FE-CLPM nur Annahmen über den Prozess der abhängigen Variablen (Yt) trifft. Allerdings führt auch das FE-CLPM zu verzerrten Schätzungen, wenn diese Annahmen verletzt sind (siehe Studie 1).

Insgesamt verdeutlichen die Befunde, dass starke Annahmen getroffen werden müssen, um Cross-Lagged Effekte kausal interpretieren zu können. Eine besondere Bevorzugung von Modellen der alternativen Ansätze (RI-CLPM, GCLPM und FE-CLPM) gegenüber dem CLPM ist aufgrund der vorliegenden Befunde nicht gerechtfertigt. Abschließend sollen Konsequenzen für die Anwendung von Cross-Lagged Panelmodellen in der Forschungspraxis diskutiert werden: (1) die begrenzte Aussagekraft des Modellfits für die Wahl des Analysemodells, (2) die Rolle des längsschnittlichen Designs und (3) die inhaltliche Interpretation des Cross-Lagged Effekts (Lüdtke & Robitzsch, 2021).

Wie stark ist die Evidenz für einen kausalen reziproken Effekt? Ein Vergleich traditioneller und neuer Verfahren zur Untersuchung des reziproken Zusammenhangs von Selbstkonzept und Leistung

Nicolas Hübner, Wolfgang Wagner, Steffen Zitzmann, Benjamin Nagengast
Universität Tübingen, Germany

Theoretischer Hintergrund. Studien, die sich genauer mit Fragen zu reziproken Entwicklungsprozessen unterschiedlicher Merkmale von Schülerinnen und Schülern (SuS) oder Lehrkräften auseinandersetzen, können dabei helfen, ein besseres Verständnis über Lehr- und Lernprozesse zu erhalten. Intensiv untersucht wurde beispielsweise das Reziproke Effektmodell (REM) der Motivation und der Leistung von SuS (Arens et al., 2017; Sewasew, Schroeders, Schiefer, Weirich & Artelt, 2018). Ein besonders prominentes Modell zur Untersuchung derartiger Zusammenhänge ist das Cross-Lagged Panel Modell (CLPM). In diesem Modell werden längsschnittliche Daten genutzt und die Schülerleistung und Motivation zu einem späteren Messzeitpunkte durch die Schülerleistung und Motivation zum vorhergehenden Messzeitpunkte zu erklären. Bei der Untersuchung der reziproken Zusammenhänge mit diesem Modell ist insbesondere der Zusammenhang zwischen der vorherigen Motivation und der späteren Leistung (unter Kontrolle der vorherigen Leistung) und der Zusammenhang zwischen der vorherigen Leistung und der späteren Motivation (unter Kontrolle der vorherigen Motivation) von Interesse (die sog. kreuzverzögerten Pfade). Zahlreiche Studien interpretieren die entsprechend gefundenen Zusammenhänge als „kausale Evidenz“ für bestimmte zugrundeliegende theoretische Modelle (Marsh & Martin, 2011;

Pinxten, Fraine, van Damme & D'Haenens, 2010). Typischerweise erfolgt diese Interpretation ohne eine Diskussion der Plausibilität zentraler Annahmen kausaler Inferenz. Allerdings wiesen Hamaker, Kuiper und Grasman (2015) und Usami, Murayama und Hamaker (2019) auf Limitationen des CLPMs zur Schätzung kausaler Effekte hin, beispielsweise weil die Annahme, dass unbeobachtete Störfaktoren keine Rolle spielen (no unobserved confounding), praktisch unplausibel erscheint. Stattdessen betonen sie die Vorteile alternativer Modelle/Methoden wie dem Random Intercept (RI) CLPM oder Gewichtungsansätzen. Lüdtke und Robitzsch (2022) konnten allerdings kürzlich in Simulationsstudien aufzeigen, dass auch das RI-CLPM in bestimmten Szenarien zu verzerrten Schätzungen führt. Ein Vergleich der Ergebnisse der alternativen und klassischen Verfahren zur Untersuchung des REMs von Motivation und Leistung ist für eine nähere Einordnung und Beurteilung bestehender Evidenzen von zentraler Bedeutung.

Fragestellung. Zu welchen Ergebnissen kommen die aktuell diskutierten Modelle/Methoden in Bezug auf die Schätzung des reziproken Zusammenhangs zwischen dem Selbstkonzept und der Leistung von SuS in Mathematik.

Methode. Es wurden Daten von 3.757 SuS der TRAIN Studie verwendet (Rose et al., 2013). In dieser Studie wurden SuS von Mittelschulen aus Sachsen und Haupt- und Realschulen aus Baden-Württemberg von Jgst. 5 bis 8 wiederholt hinsichtlich ihrer Motivation und Leistung untersucht. Zur Beantwortung der Fragestellung wurden CLPM, Full-Forward (FF) CLPM und RI-CLPM spezifiziert. Darüber hinaus erfolgte die Schätzung der reziproken Zusammenhänge zwischen dem Selbstkonzept und der Leistung unter Verwendung von zwei Gewichtungsprozeduren für kontinuierliche Treatmentvariablen (Covariate Balanced Generalized Propensity Score Weighting und Entropy Balancing; Fong, Hazlett & Imai, 2018; Tübbicke, 2021). Bei der Spezifikation der Gewichtungsansätze wurden zahlreiche Kovariaten berücksichtigt (z.B. Variablen in Bezug auf die Demographie, den sozialen und kulturellen Hintergrund, Leistungsmerkmale und die Persönlichkeit). Die CLPM, FF-CLPM und RI-CLPM wurden mit und ohne ein identisches Set von Kovariaten spezifiziert. Fehlende Werte wurden multipel imputiert. Darüber hinaus wurden clusterrobuste Standardfehler zur Berücksichtigung der Nestung von Schülerinnen und Schülern in Schulen berechnet.

Ergebnisse und ihre Bedeutung. Die Ergebnisse der Studie legen nahe, dass CLPM und FF-CLPM (mit und ohne Kovariaten) zu sehr vergleichbaren Ergebnissen gelangen wie die Gewichtungsverfahren: Hier findet sich Evidenz im Sinne reziproker Zusammenhänge zwischen der Leistung und dem Selbstkonzept. Die Befunde des RI-CLPM hingegen waren unsystematisch. Die Ergebnisse der vorliegenden Studie ähneln damit insgesamt Ergebnissen anderer Studien, die teilweise deutliche Unterschiede zwischen den Schätzungen auf Basis von RI-CLPM und CLPM/FF-CLPM fanden (z.B. Bailey, Oh, Farkas, Morgan & Hillemeier, 2020; Ehm, Hasselhorn & Schmiedek, 2019). Zusammenfassend scheinen klassische Modelle (ohne Kovariaten) in der vorliegenden Studie zu relativ ähnlichen Ergebnissen zu gelangen wie Verfahren, bei denen die Erfüllung bestimmter Annahmen, die zur Schätzung kausaler Effekte erforderlich sind, insgesamt plausibler erscheinen.

Beeinflussen Schulreformen das Lernverhalten an der Universität? Ein Vergleich traditioneller und neuer Varianten der Difference-in-Differences-Schätzung

Jakob Schwerter[1], **Nicolai Netz**[2], **Nicolas Hübner**[3]
[1]Technische Universität Dortmund, [2]Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), [3]Universität Tübingen, Germany

Theoretischer Hintergrund. Bildungsreformen bieten Politiker:innen Möglichkeiten zur nachhaltigen Verbesserung von Bildungssystemen. Wie Daten des OECD Education Policy Outlook zeigen, vergeht kaum ein Jahr, in dem sich das internationale Bildungswesen nicht mit umfangreichen reforminduzierten Veränderungen konfrontiert sieht. Zur Beantwortung der Frage nach der Wirksamkeit derartiger Reformen ist eine systematische Begleitforschung von zentraler Bedeutung. Eine insbesondere in der Ökonomik populäre Methode zur Reformanalyse ist die Difference-in-Differences (DiD) Analyse (Angrist & Pischke, 2008). Bei diesem quasi-experimentellen Verfahren wird der Effekt einer Reform, die beispielsweise in einem Bundesland eingeführt wurde, berechnet, indem die Entwicklung des reformierten Bundeslandes mit derjenigen eines nicht-reformierten Bundeslandes verglichen wird. Ein überwiegender Teil der entsprechenden Reformliteratur verwendet zur Berechnung des Reformeffekts eine Two-Way Fixed Effects (TWFE) Regression. Neuere Literatur weist allerdings auf Verzerrungen bei dieser Schätzmethode hin, wenn Reformen gestaffelt eingeführt werden, d.h. zu unterschiedlichen Zeitpunkten (Roth et al., 2022). Ein zentrales Problem der bisher verwendeten TWFE-Regression liegt darin, dass bei einer gestaffelten Reformeinführung die Entwicklung der reformierten Bundesländer mit der Entwicklung jener Bundesländer verglichen wird, die die Reform bereits implementiert haben. Dies kann zu einer substantiellen Verzerrung der Koeffizientenschätzungen in der TWFE-Regression führen (Roth et al., 2022). Um diese Herausforderung der TWFE-Regressionen zu adressieren, wurden kürzlich verschiedene Lösungsvorschläge präsentiert, von denen die Schätzung des gewichteten gruppenzeitlichen durchschnittlichen Reformeffekts besonders prominent diskutiert wird (Callaway & Sant'Anna, 2021; DiD-CA). Dieser Ansatz berechnet den Reformunterschied zunächst separat für jeden Erhebungspunkt, bevor die ermittelten Unterschiede anschließend aggregiert werden. Somit können u.a. bei späteren Reformeinführungen gezielt Vergleiche mit Bundesländern ausgeschlossen werden, die die Reform bereits umgesetzt haben. Ein Vergleich der klassischen und neuen Möglichkeiten der Reformeffektschätzung ist im Kontext deutscher gestaffelter Reformen (z.B. der G8 Reform oder der Oberstufenreform) bisher noch nicht erfolgt.

Fragestellung. Finden sich bei Verwendung des DiD-CA und der TWFE-Schätzung unterschiedliche G8-Reformeffekte in Bezug auf eine Veränderung des Lernaufwands von Studierenden? Die G8-Reform verringerte die Schulzeit an Gymnasien in 13 Bundesländern von 9 auf 8 Jahre. Hierzu wurde die durchschnittliche wöchentliche Stundenanzahl von 30- 32 auf 32-34 Stunden pro Woche erhöht (Huebener et al., 2017). Verschiedene Studien legen nahe, dass Schüler:innen aus G8-Jahrgängen insgesamt mehr stressbezogene gesundheitliche Probleme sowie am Ende der Sekundarstufe II zum Teil geringere Kompetenzen aufweisen als Schüler:innen aus G9-Jahrgängen (Hübner et al., 2022; Huebener & Marcus, 2017; Marcus et al., 2020; Quis, 2018). Bisher existieren allerdings kaum Studien, die sich näher mit den langfristigen Auswirkungen der G8-Reform auseinandersetzen, beispielsweise mit lernbezogenem Verhalten im Studium. Die Betrachtung beider Methoden ist hierbei

wichtig, um aufzuzeigen, welchen Einfluss die gestaffelte Reformeinführung auf das Schätzergebnis der TWFE-Regression haben kann.

Methode. Zur Beantwortung der Forschungsfrage wurden Daten der deutschen Studierenden-Sozialerhebung des DZHW (Apolinarski et al., 2021) der Jahre 2003 bis 2016 verwendet. Diese umfassen Daten von Studierenden der Abiturjahrgänge 1998 bis 2016. Der gepoolte Datensatz umfasst insgesamt N = 71.426 Studierende (59 % weiblich, MAlter = 23).

Für ein erfolgreiches Studium ist insbesondere der Lernaufwand von zentraler Bedeutung (Andrietti & Velasco, 2015; Schwerter et al., 2022). Dieser wurde mithilfe der Selbsteinschätzungen in Stunden pro Woche für Vorlesungen, das Selbststudium und die aufgewendete Arbeit erfasst.

Ergebnisse und ihre Bedeutung. Unter Verwendung der DiD-CA Methode zeigen sich negative Auswirkungen der Reform auf den Lernaufwand. Bei der TWFE-Schätzung ist der Reformeffekt hingegen nicht statistisch signifikant. Die Verwendung der TWFE-Schätzung würde somit fälschlicherweise suggerieren, dass die veränderte Lernzeit in der Sekundarschulzeit keinen Einfluss auf den Lernaufwand im Studium hätte. Dieser Unterschied legt die Notwendigkeit einer Untersuchung gestaffelter Reformen mit Verfahren nahe, die diese Staffelung berücksichtigen. Die mittels DiD-CA Methode gefunden Reformeffekte liefern darüber hinaus eine Erklärung für bereits berichtete Leistungsunterschiede von G8-Studierenden (Doersam & Lauber, 2019).

Transparenz und Reproduzierbarkeit von Meta-Analysen zu Self-Beliefs von Schülerinnen und Schülern: Ein Meta-Review

Thorben Jansen[1], Jennifer Meyer[1], Jens Möller[2]
[1]Leibniz Institute for Science and Mathematics Education at Kiel University, [2]Institute for Psychology of Learning and Instruction, Kiel University, Germany

Hintergrund

Meta-Analysen sind ein wichtiges Instrument, um empirische Forschungsergebnisse zusammenzufassen. Ihre Methodik sollte transparent beschrieben und reproduzierbar sein, um sicherzustellen, dass die Datengrundlagen der Interpretationen von hoher Qualität sind und diese auf neue Studien und Auswertungsmethoden angepasst werden können. Zu diesem Ziel sollten Meta-Analysen Richtlinien wie z.B. die Preferred Reporting Items for Systematic Reviews und Meta-Analysen (PRISMA; Page et al., 2020) befolgen. Polanin et al. (2020) zeigen, dass psychologische Meta-Analysen verschiedener Fachrichtung jedoch nur knapp mehr als die Hälfte der Qualitätskriterien erfüllen, insbesondere wurden selten die Primärdaten und der statistische Code angegeben. Unklar ist bislang, inwieweit Meta-Analysen aus dem Bereich der Bildungsforschung die Qualitätskriterien etabliert haben. Eine empirische Evaluierung der Transparenz von Meta-Analysen kann Hinweise auf methodische Limitationen bisheriger meta-analytischer Forschung geben und so dazu beitragen, dass diese Praktiken für künftige Autor*innenteams, Geldgeber und die Redaktionen von Zeitschriften zur Routine werden. Wir untersuchen die Transparenz und Reproduzierbarkeit innerhalb einer Sekundäranalyse von Meta-Analysen aus dem Forschungsfeld der Self-Beliefs (Selbstkonzept, Selbstwirksamkeit und Selbstwert) von Schülerinnen und Schülern.

Fragestellungen

Wie transparent und reproduzierbar sind Meta-Analysen zu den Self-Beliefs von Schülerinnen und Schülern? Welche Qualitätskriterien werden am häufigsten und wenigsten erfüllt? Erfüllen neuere Meta-Analysen mehr Kriterien der Transparenz und Reproduzierbarkeit?

Methode, Daten und Ergebnisse

Die Transparenz und Reproduzierbarkeit wurde innerhalb einer Sekundäranalyse von 94 Meta-Analysen untersucht. Jede Meta-Analyse berichtete mindestens eine Effektstärke über den Zusammenhang einer Variable (z.B. Schulleistungen, Interventionen) und den Self-Beliefs von Schülerinnen und Schülern im Alter zwischen Kindergarten und Klassenstufe 12. Von den Meta-Analysen wurden 20 nach 2019, 27 zwischen 2019 und 2010, 25 zwischen 2009 und 2000, und 25 vor 2000 veröffentlicht. Für jede Meta-Analyse kodierten wir jeweils die 29 von Polanin et al. (2020) vorgeschlagenen Kriterien. Durchschnittlich wurden 54% der Kriterien erreicht. Die Kriterien wurden in neun Kategorien zusammengefasst, diese sind:

• Funding und Präregistration: Meta-Analyse wurde präregistriert und gibt Informationen zum Funding (16 % erfüllt)

• Publikationsbias Analyse: Meta-Analyse führte eine Publikationsbasis Analyse (z.b. funnel plots) aus (44 % erfüllt)

• Ein/Ausschlusskriterien: Meta-Analyse beschreibt, welche Population, Design, Prädiktoren, Outcomes und Publikationsstatus eingeschlossene Studie haben (86 % erfüllt)

• Effektstärkenberechnung: Meta-Analyse beschreibt, wie die Effektstärken transformiert, abhängige Effektstärken behandelt und mittlere Effektstärken berechnet wurden (82 % erfüllt)

• Primärdaten: Die Primärdaten zur Berechnung der mittleren Effektstärken sind zugänglich (12 % erfüllt)

• Kodierprozess: Das vollständige Kodierhandbuch ist öffentlich zugänglich (30 % erfüllt)

• Suchstrategie: Meta-Analyse beschreibt ihre Suchparamenter, das Datum der Suche und die verwendeten Datenbanken (51 % erfüllt)

• Transparente Literatursuche: Meta-Analyse beschreibt, wie viele Studien an welchem Schritt der Suche ausgeschlossen wurden, z.B. in einem Flow Diagramm (36 % erfüllt)

• Daten für Moderatoranalysen: Meta-Analyse veröffentlicht die Primärdaten der Moderatoranalysen (36 % erfüllt)

Bedeutung

Unsere Befunde zeigen, dass die Meta-Analysen im Forschungsfeld der Self-Beliefs etwas mehr als die Hälfte der kodierten Kriterien erfüllen und finden eine positive zeitliche Entwicklung: Alle Kriterien werden mit jüngerem Veröffentlichungsjahr häufiger erfüllt, jedoch auch die neuesten Meta-Analysen nicht alle Kriterien. Die Meta-Analysen hatten in hohem Maße transparente Ein/Ausschlusskriterien und berichteten den Prozess der Ermittlung der mittleren Effektstärken. Entwicklungspotenzial gibt es besonders bei der Veröffentlichung der Primärdaten, Veröffentlichung des statistischen Codes und die Präregistrierung der Meta-Analysen. Dieser Beitrag impliziert für zukünftige Meta-Analysen zu den

Self-Beliefs, dass Autorinnen, Autoren und Zeitschriften besonderen Wert auf die Veröffentlichung der Primärdaten und des statistischen Codes legen sollten, da diese Kriterien am wenigsten erfüllt sind und die auch die Transparenz von anderen Kriterien erleichtern würde.

F-21: Qualitätsmerkmale und Wirksamkeit von Erklärvideos

Zeit: Donnerstag, 02.03.2023: 10:30 - 12:15 · *Ort:* R14 R02 B07

Qualitätsmerkmale und Wirksamkeit von Erklärvideos

Chair(s): **Ann-Kathrin Jaekel** (Hector-Institut für Empirische Bildungsforschung, Universität Tübingen), **Malte Ring** (Eberhard Karls Universität Tübingen, Lehrstuhl für Ökonomische Bildung und Wirtschaftsdidaktik)
*Diskutant*in(nen):* **Katharina Scheiter** (Universität Potsdam)

Erklärvideos sind ein Medium zur Wissensvermittlung und -aneignung, das über ein großes Potenzial verfügt. Die NutzerInnen können damit in eigenem Tempo Sachverhalte beliebig oft wiederholen, die ErstellerInnen können selbst komplexe und aufwändige Lerninhalte anschaulich darstellen. Aktuelle Studien zur Nutzung von Erklärvideos zeigen, dass diese ein wesentlicher Bestandteil der Alltagsrealität von Kindern, Jugendlichen und jungen Erwachsenen sind (Feierabend et al., 2020a; 2020b). Hinsichtlich ihrer Wirkungsweise konnten empirische Studien zeigen, dass sich Erklärvideos positiv auf die Lernleistung, die Aufmerksamkeit sowie die Motivation der NutzerInnen auswirken können (Jill et al., 2019; van der Meij & van der Meij, 2014; Lloyd & Robertson, 2012; Sargent et al., 2011).

Die bloße Nutzung und Bereitstellung von Erklärvideos reichen dabei jedoch nicht aus. Vielmehr muss deren Einsatz im Bildungskontext stets im Hinblick auf ihre Qualität sowie ihren didaktischen Mehrwert betrachtet werden. So sind derzeit noch viele Fragen zur Gestaltung und Wirkung von Erklärvideos offen. Beispielsweise zeigen sich zur Sichtbarkeit der Lehrkraft oft unterschiedliche Effekte (Henderson & Schroeder, 2021). Auch ist unklar, welche Elemente eines Erklärvideos besonders förderlich für den Kompetenzgewinn sind (REFs). In diesem Symposium befassen wir uns daher mit der Qualität und der Nutzung von Erklärvideos im Schul- und Hochschulkontext aus verschiedenen Perspektiven: In zwei fachunabhängigen Beiträgen gehen wir der Frage nach, welche Qualitätsmerkmale bei der Erstellung von Erklärvideos besonders wichtig sind. Zwei weitere, fachdidaktische Beiträge beschäftigen sich mit der Frage, welchen Mehrwert Erklärvideos in den jeweiligen Fächern im Hinblick auf traditionellere Lernmethoden sowie dem Zuwachs an inhaltlichem Wissen bieten.

Beitrag 1 stellt einen fachunabhängigen Ratingbogen zur Erfassung der Qualität von Erklärvideos vor. Erprobt an Videos die a) durch Lehrkräfte selbst erstellt wurden und b) professionellen Ursprungs der Plattform Youtube entnommen sind, wird der aktuelle Stand der Validierung des Ratingbogens vorgestellt.

Beitrag 2 geht anhand der Nutzung von Eye-Tracking-Daten der Frage nach, wie sich die Sichtbarkeit einer Lehrkraft in Erklärvideos auf die visuelle Aufmerksamkeit der NutzerInnen sowie deren Lernergebnisse auswirkt.

Beitrag 3 beschäftigt sich aus Sicht der Physikdidaktik mit der Frage, ob Erklärvideos mit alltagsnahen Beispielen zu verzerrten Vorstellungen des fachlichen Verständnisses, sogenannten Verstehensillusionen führen können. Im Hinblick auf diese Frage schließt sich eine Gegenüberstellung der Erklärvideos mit Textmaterialien an.

Beitrag 4 untersucht, inwiefern sich durch einen schrittweisen Aufbau von Visualisierungen in Erklärvideos ein Zuwachs des Verständnisses wirtschaftlicher Graphen bei Studierenden verzeichnen lässt.

Beiträge des Symposiums

Ein fächerübergreifendes Instrument zur Erfassung der Qualität von Erklärvideos im schulischen Kontext

Ann-Kathrin Jaekel, Richard Göllner
Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

Egal ob zur Wiederholung, im Flipped Classroom oder zum Umgang mit einer heterogenen Schülerschaft: Erklärvideos sind eine vielseitig einsetzbare Methode für jedes Unterrichtsfach. Zwar haben sich während der COVID 19-Pandemie insbesondere Erklärvideos, die von Lehrkräften selbst erstellt waren, als besonders positiv herausgestellt und nicht solche, die von externen Plattformen stammten (Jaekel et al., 2021). Jedoch bieten auch Plattformen wie Youtube oder Planet Schule eine immense Anzahl hoch professionell erstellter Videos zu nahezu jedem Themenbereich, welche Lehrkräfte für ihren Unterricht nutzen können.

Hinsichtlich der Qualität von Erklärvideos haben sich in den vergangenen Jahren insbesondere die Forschung der Fachdidaktiken oder aus dem Hochschulkontext mit relevanten Qualitätsmerkmalen von Erklärvideos beschäftigt. Diese identifizierten beispielsweise eine klare Struktur, zielgerichtete Erklärungen, Visualisierungen und das Anknüpfen an das Vorwissen der Schülerinnen und Schüler als wichtige Merkmale von für das Lernen der Schülerinnen und Schüler (Brame, 2016; Kay, 2014; Kulgemeyer 2018; Schopf, 2020). Diese Merkmale sind auch aus der Forschung zu regulärem schulischem Unterricht bekannt (Kunter & Traut-wein, 2013), sodass sich beide Forschungsfelder als Ausgangspunkt für die Identifikation von fächerübergreifenden Qualitätsmerkmalen von Erklärvideos dienen können.

Insgesamt beschränkt sich die Forschung auf spezielle Aspekte von Unterricht wie einem bestimmten Schulfach oder den Hochschulkontext, sodass bislang keine fächerübergreifen-den Qualitätsmerkmale für den schulischen Kontext identifiziert werden konnten. Daher gehen wir in der vorliegenden Studie den folgenden Forschungsfragen nach:

1) Welche Qualitätsmerkmale lassen sich aus der bisherigen Forschung zu Erklärvideos sowie aus der Unterrichtsqualitätsforschung auf Erklärvideos für den schulischen Unterricht übertragen?

2) Inwiefern lassen sich diese Qualitätsmerkmale für von Lehrkräften selbst erstellte Lernvideos sowie für professionell erstellte Youtube-Videos empirisch bestätigen?

3) Zeigen sich Unterschiede zwischen der Übereinstimmung von unterschiedlichen Ratern zwischen selbst erstellten Erklärvideos von Lehrkräften und Youtube-Videos?

Zur Beantwortung der Fragestellung haben wir 22 von Lehrkräften selbst erstellte Videos ge-nutzt, wobei keine Einschränkung der Fächer oder der Schulart vorlag. Für die professionell erstellten Youtube-Videos haben wir zum Themenbereich „Kommasetzung“ für das Fach Deutsch sowie zum Themenbereich „Bruchrechnen“ für das Fach Mathematik jeweils zehn Videos verwendet. Die Videos haben eine Länge zwischen 2:58 min – 13:52 min. Alle Videos wurden von zwei Raterinnen unabhängig voneinander entlang des Ratingmanuals bewertet. Auf dieser Grundlage wurden die Ratings anhand einer Generalisierbarkeitsstudie ausgewertet (Brennan, 2001). Als Facetten dienten Rater, Lernvideos und Items, wobei Lernvideos innerhalb von Ratern und Items innerhalb von

Lernvideos geschachtelt waren. Da beide Rater alle Videos beurteilten und alle Lernvideos anhand der gleichen Items beurteilt wurden, handelte es sich um ein balanciertes Untersuchungsdesign. Die von Lehrkräften selbsterstellten Lernvideos und die Youtube-Videos wurden getrennt voneinander analysiert.

Die Ergebnisse der Generalisierbarkeitsstudie zeigten, dass mittels der Facetten ein Großteil der Gesamtvarianz für Videos beider Quellen erklärt werden konnte (selbsterstellte Lernvideos: 78%; Youtube-Videos: 79%). Die größten Varianzanteile ergaben sich dabei für Unter-schiede zwischen Items (selbsterstellte Lernvideos: 46%; Youtube-Videos: 43%). Demgegen-über erklärten die Lernvideos (selbsterstellte Lernvideos: 0%; Youtube-Videos: 4%) und die Rater (selbsterstellte Lernvideos: 0%; Youtube-Videos: 1%) nur einen unbedeutenden Anteil der Gesamtvarianz in den Ratings. Zudem zeigte die Interaktion zwischen Lernvideos und Items einen substanziellen Varianzanteil (selbsterstellte Lernvideos: 27%; Youtube-Videos: 16%). Unterschiede der Ratings zwischen Unterrichtsvideos waren entsprechend auf die Spezifität der Items bzw. deren Kopplung mit den Lernvideos zurückzuführen. In diesem Zusammenhang erwiesen sich beispielsweise die Items „Die Lehrkraft fasst den Lerninhalt am Ende des Videos zusammen.“ und „Die Lehrkraft stellt die Relevanz des Themas für die Lernenden heraus.“ als besonders geeignet, um Unterschiede zwischen Videos abzubilden.

Abgelenkt durch einen Talking Head? Effekte einer sichtbaren Lehrperson in Lernvideos

Christina Sondermann[1], Markus Huff[2], Martin Merkt[1]
[1]Deutsches Institut für Erwachsenenbildung, Bonn, [2]Eberhard Karls Universität Tübingen, Tübingen & Leibniz-Institut für Wissensmedien, Tübingen

Theoretischer Hintergrund

Sollte eine Lehrperson in Lernvideos neben den eigentlichen Lerninhalten eingeblendet werden? Unterschiedliche Theorien legen diesbezüglich sowohl Vor- als auch Nachteile nahe: Einerseits nimmt die Social-Agency-Theorie (Mayer, 2014) an, dass soziale Hinweisreize im Lernmaterial soziale Reaktionen auslösen können, die wiederum zu einer tieferen Verarbeitung und verbessertem Lernen führen. Andererseits kann eine sichtbare Lehrperson nach der Cognitive-Load-Theory (Sweller et al., 2011) und der Cognitive Theory of Multimedia Learning (Mayer, 2014) zu Ablenkung, geteilter Aufmerksamkeit und kognitiver Überlastung führen.

Auch die empirische Evidenz zeigt heterogene Befunde, wobei Studien teilweise positive und teilweise ausbleibende Effekte der sichtbaren Lehrperson auf Lernergebnisse feststellen konnten (Henderson & Schroeder, 2021). In einer aktuellen Onlinestudie (Autoren, 2022) fanden wir hingegen negative Effekte eines Talking Heads (TH) auf die Lernergebnisse. Da in letzterer Studie die Blickbewegungen der Lernenden nicht aufgezeichnet werden konnten, blieb offen, ob sich die Lerneinbußen durch vermehrte visuelle Aufmerksamkeit auf dem TH erklären lassen. Bisherige Eyetracking-Studien haben bereits gezeigt, dass die eigentlichen Lerninhalte seltener angesehen wurden, wenn daneben eine Lehrperson sichtbar war (Colliot & Jamet, 2018). Hinsichtlich der Bewertung des Lernmaterials zeigten viele Studien positive Effekte einer sichtbaren Lehrperson (Autoren, 2022; Wang & Antonenko, 2017).

Fragestellung

In diesem Eyetracking-Experiment wurde untersucht, ob der lernhinderliche Effekt des TH aus der vorangehenden Studie (Autoren, 2022) repliziert und durch reduzierte visuelle Aufmerksamkeit auf dem Lerninhalt erklärt werden kann. Basierend auf der vorangehenden Studie haben wir geringere Lernergebnisse und höhere Bewertungen (z.B. Zufriedenheit) für Videos mit TH als für Videos ohne TH erwartet. Außerdem haben wir angenommen, dass die visuelle Aufmerksamkeit (Fixierungsdauer) für die Lerninhalte geringer ist, wenn ein TH sichtbar ist. Weiterhin sollte im Rahmen einer explorativen Forschungsfrage untersucht werden, ob die visuelle Aufmerksamkeit auf den Lerninhalten den TH-Effekt auf die Lernergebnisse moderiert. Sämtliche Hypothesen und Forschungsfragen wurden präregistriert (https://osf.io/rcde6/?view_only=fb6aa775e1b14d30826d2af29071bc51).

Methode

Wie präregistriert sahen 96 Studierende (73 weiblich, MAlter=22.65 Jahre, SDAlter=3.77) acht kurze Videos (Dauer ca. 1 Minute) zu acht verschiedenen Themen (z.B. Mönchsgrasmücke). Mithilfe eines Within-Designs wurde variiert, ob in den Videos ein TH zu sehen war oder nicht, wobei die Tonspur komplett identisch war. Dabei wurde balanciert, welches Videothema mit und ohne TH gezeigt wurde und ob der TH in der rechten oder linken unteren Ecke zu sehen war. Die Videos bestanden aus Folien mit grafischen Inhalten, die nach und nach erschienen.

Nach einer anfänglichen Selbstauskunft zu Vorwissen und Interesse bzgl. der acht Videothemen und einer Kalibrierung des Eyetrackers folgten die Lernvideos. Währenddessen wurden die Blickbewegungen der Lernenden aufgezeichnet (Tobii Pro Nano, 60Hz). Nach jedem Video wurden verschiedene Bewertungsmaße (z.B. Zufriedenheit) erfragt. Anschließend füllten die Lernenden den Wissenstest aus, der aus 24 offenen Fragen (drei pro Thema) bestand und Fakten aus den Videos (einzelne Wörter oder Zahlen) abfragte. Darauf folgte eine Picture Recognition Task, die aus 24 Fragen (drei pro Thema) bestand. Dabei sollten die Lernenden aus vier Bildern dasjenige auswählen, das im jeweiligen Lernvideo zu sehen war. Abschließend folgten demografische Fragen und Kontrollfragen.

Ergebnisse und ihre Bedeutung

Im Zuge vorläufiger Analysen zeigte die messwiederholte ANOVA keine Effekte des TH auf den Wissenstest und die Picture Recognition Task, beide $F<1$. Die Analysen zu Eyetrackingdaten und Bewertungsmaßen stehen noch aus, werden aber im Rahmen des Symposiums vorgestellt. Die ausstehenden Analysen beleuchten, ob der Einfluss des TH auf die Lernmaße durch die visuelle Aufmerksamkeit auf den Lerninhalten moderiert wird. Bei Betrachtung der vorläufigen Ergebnisse liefert unsere Studie weitere Hinweise darauf, dass eine sichtbare Lehrperson in Lernvideos häufig keine Effekte auf Lernmaße hat (vgl. Henderson & Schroeder, 2021). Ein möglicher Grund für die von der vorangehenden Studie abweichenden Lernergebnisse könnte das veränderte Studienformat (Laborexperiment statt Online-Experiment) sein, weil der formalere Charakter eines Laborexperiments lernhinderliche TH-Effekte aufgehoben haben könnte.

Fördern Erklärvideo Wissen – oder eine Verstehensillusion? Zwei Experimente aus der Physik

Christoph Kulgemeyer
Universität Bremen

Theoretischer Hintergrund

Erklärvideos sind zu allen Themen der Physik-Schulcurricula mittlerweile frei erhältlich (Kulgemeyer & Peters, 2016). Es gibt eine breite Forschung an Kriterien für die Instruktionsqualität von Erklärvideos (z.B. Kulgemeyer, 2018; Findeisen, Horn & Seifried, 2019; Mayer, Fiorella & Stull, 2020). Als alternatives Instruktionsmaterial kommen Lehrbuchtexte infrage. Die Studienlage zum Vergleich der Medien ist jedoch nicht eindeutig. Lloyd und Robertson (2012) identifizieren Videos als überlegen bezüglich des Wissenserwerbs, Lee und List (2019) das Lernen aus Texten, List und Ballenger (2019) finden keinen Unterschied.

Erklärvideos werden häufig auf Videoplattformen konsumiert. Allerdings bieten auch große Kanäle aus fachdidaktischer Sicht problematische Inhalte an, z.B. Videos, die scheinbar alltagsnahe Erklärungen bieten, die aber fachlich falsch sind und sogar bekannte Lernhindernisse darstellen (Schülervorstellungen) (Kulgemeyer & Wittwer, 2022). Es könnte sein, dass Lernende diese wegen der scheinbaren Alltagsnähe aber für besser verständlich halten.

Die Überzeugung, eine Erklärung sei gut und führe zu so hohem Verstehen, dass Folgeunterricht nicht mehr notwendig sei, wird als „Verstehensillusion“ bezeichnet (Wittwer & Renkl, 2008; Paik & Schraw, 2013). Einige Studien legen nahe, dass Videos Verstehensillusionen befördern (z.B. Lowe, 2003). Zuletzt wurde aber mit der „Shallowing Hypothese“ (Salmerón et al. 2020) die Wirkung des Mediums genauer beschrieben: Online-Medien kommen vor allem im privaten Bereich mit oberflächlicher Interaktion vor, die sich dann auf Lernkontexte überträgt. Umgekehrt kommen Lehrbuchtexte in formalen Lernkontexten vor und suggerieren so Seriosität, was den eigenen Lernerfolg überschätzen ließe. Es ist also weder klar, ob Videos im Vergleich zu geschriebenen Erklärungen zu einer höheren Verstehensillusion führen, noch ob Schülervorstellungen als scheinbar einfache Alternativerklärungen von Novizen akzeptiert werden und Verstehensillusionen hervorrufen.

Fragestellung

Die Studie besteht aus zwei Experimenten zu Wissenserwerb und Verstehensillusion in der Physik: (1) Ein Vergleich fachlich anschlussfähiger Videos mit Videos, die alltagsnahe Schülervorstellungen als falsche Alternativerklärungen anbieten, (2) Ein Vergleich von Lehrbuchtexten und Erklärvideos.

Methode

Es wurden zwei Erklärvideos zur Einführung in den Kraftbegriff entwickelt; eines fachlich anschlussfähig, eines nach Vorbild eines YouTube-Videos auf Basis alltagsnaher Schülervorstellungen. Beide Videos sind sonst vergleichbar, u.a. nach Kriterien der Instruktionsqualität. Die Videos wurden randomisiert zwei Gruppen Lernender mit geringem Vorwissen zugeteilt („Schülervorstellungsgruppe“: N = 69, Gruppe „fachlich korrekt“: N = 80). Gemessen wurden u.a. deklaratives Wissen zum Kraftbegriff (Multiple-Choice, 11 Items, α=0,71), das fachlich falsche „Schülervorstellungswissen“ (d.h. die „Lerngelegenheit“ des falschen Videos) (Multiple-Choice, 11 Items, α=0,85), die Verstehensillusion (Fünfstufige Likert-Skala, z.B. „Ich habe verstanden, was Kraft in der Physik ist“, 10 Items, α=0,75) und Kontrollvariablen. Im zweiten

Experiment wurde eine vergleichbare Gruppe randomisiert auf die Bedingungen „fachlich korrektes Video“ und „Transskript des Videos als Lehrbuchtext aufbereitet“ aufgeteilt (N_Video = 78; N_Text: 72). Gemessen wurden dieselben Variablen. Die Unterschiede der Besetzung der Bedingungen resultierten aus mangelnder Zustimmung zur Datennutzung.

Ergebnisse und ihre Bedeutung

Die Gruppen „Schülervorstellung“ und „fachlich korrekt“ überschätzen beide ihr Verständnis. Die Gruppen „Schülervorstellung“ lernt mehr falsches Schülervorstellungswissen als die andere Gruppe fachlich korrektes Wissen ($F(1,147) = 18.39$, $p < 0.001$, $\eta 2 = 0.11$). Lernende aus der Gruppe „Text“ lernen ebenso viel über das Kraftkonzept wie Lernende aus der Gruppe „Video“, haben aber eine geringfügig stärkere Verstehensillusion ($F(1,147) = 6.65$, $p = 0.011$, partial $\eta 2 = 0.043$; Verstehensillusion mit Wissen post kontrolliert).

Es wird die Hypothese abgeleitet, dass Lernende bei beiden Instruktionsmaterialien eine Einbettung in eine Lernumgebung benötigen, die das Verstehen infrage stellt (z.B. weiterführende Lernaufgaben). Ansonsten könnte die Verstehensillusion dazu führen, dass in Folgeunterricht keine hohe kognitive Aktivierung hergestellt werden kann. Es könnte zudem sein, dass Lernende Videos mit alltagsnahen Schülervorstellungen als Alternativerklärungen bei Plattformen wie YouTube besser beurteilen, da sie aus diesen mehr (aber Falsches) lernen und eine Verstehensillusion aufweisen. Dies sollte weiter untersucht werden.

Achse für Achse – Fördert der schrittweise Aufbau von wirtschaftlichen Graphen in Erklärvideos deren vertieftes Verständnis?

Malte Ring, Afroditi Kourtoglou, Taiga Brahm
Eberhard Karls Universität Tübingen, Lehrstuhl für Ökonomische Bildung und Wirtschaftsdidaktik

Theoretischer Hintergrund

Erklärvideos sind ein beliebtes Medium und werden in unterschiedlichen Fachkon-texten zur Vermittlung von Inhalten eingesetzt (Findeisen et al., 2019). Aufgrund der ausgeprägten Verfügbar- und Zugänglichkeit bieten Erklärvideos einerseits ein großes Lernpotential (Expósito et al., 2020). Andererseits zeigen Studien, dass Vi-deos oftmals auch nicht zum gewünschten Lernergebnis führen, was mitunter mit dem Design der Videos begründet wird (Fiorella & Mayer, 2018).

Dabei kann ein schrittweiser Aufbau der Visualisierung Lernenden helfen, sich an die Inhalte des Videos besser zu erinnern (Fiorella et al., 2019; Fiorella & Mayer, 2016) – dagegen zeigt sich teilweise kein Unterschied im Hinblick auf komplexere Transferaufgaben (Fiorella et al., 2019). Die positive Wirkung wird mit unterschiedli-chen Mechanismen erklärt: Erstens lenkt der schrittweise Aufbau die Aufmerksam-keit direkt auf die relevanten Teile und unterstützt die Lernenden dabei, die Informa-tionen aus dem gesprochenen Text den richtigen Teilen der Visualisierung zuzu-ordnen (Signaling bzw. Temporal Contiguity, Überblick in Ginns, 2006; auch Noetel et al., 2022). Weiterhin führt ein schrittweiser Aufbau dazu, dass die Inhalte in klei-nere Abschnitte geteilt werden, was insbesondere Lernende mit wenig Vorwissen beim Verständnis unterstützen kann (Spanjers et al., 2010).

Noch offen ist, wie sich der schrittweise Aufbau auf das Verständnis von Graphen auswirkt. Einerseits könnte ein schrittweiser Aufbau zusätzlich dazu führen, dass die zu Grunde liegenden „Leseregeln“

und die Beziehung der einzelnen Graphen-Elemente besser verstanden werden und dadurch die Entwicklung eines genaue-ren mentalen Modells gefördert wird. Andererseits ist es auch möglich, dass eine Erklärung allein über ein Video– ohne die Möglichkeit die gelernten Prinzipien in einem komplexeren Lehr-Lern-Setting anzuwenden – nicht ausreicht, um den Um-gang mit den Graphen zu verbessern.

Fragestellung

Im Rahmen dieser Studie soll untersucht werden, inwiefern der schrittweise Aufbau der Visualisierungen in Erklärvideos ein vertieftes Verständnis wirtschaftlicher Gra-phen fördert. Auf Basis bisheriger Forschung wird angenommen, dass sich der schrittweise Aufbau positiv auf die Erinnerungsleistung bezüglich der Videoinhalte auswirkt. Ob sich auch für die komplexere Diagrammkompetenz eine positive Wir-kung einstellt, ist noch offen.

Methode und Material

Auf Basis der bisherigen Erkenntnisse wurde in einer a-priori Power-analyse eine notwendige Stichprobengröße von 100 Teilnehmenden pro Gruppe berechnet (t-Test, Cohen's d von 0,4, Power: 1 - β= 0,8) und gemeinsam mit den Hypothesen prä-registriert (https://osf.io/4zpbg/?view_only=967db03554604f7aa52ebcfb7c0f87d2). Seit Studienbeginn am 19.09.2022 haben bereits 74 Studierende teilgenommen, die nicht Wirtschaft studieren.

Es wurden zwei etwa 10-minütige Erklärvideo zum Thema Preisbildung entwickelt. Die Videos bestehen aus mehreren Abschnitten, in denen die Grundsätze des Preis-Mengen-Diagramms erklärt werden. Die Videos unterscheiden sich nur im Hinblick auf den Aufbau der Visualisierung: In einem Video werden die Graphen Schritt für Schritt – jeweils passend zur verbalen Erklärung aufgebaut. Im anderen Video wird die gesamte Visualisierung bereits zu Beginn der Erklärung gezeigt. Im Anschluss wird die Lernleistung mit verschiedenen Skalen erfasst. Die Erinne-rungsleistung wird mit fünf Items erhoben, die sich auf die Inhalte des Videos be-ziehen. Verschiedene Facetten von Diagrammkompetenz (Lesen, Interpretieren, Auswählen, Anwenden, Fehler identifizieren und evaluieren) werden mithilfe eines Diagrammkompetenztests erfasst. Zusätzlich werden allgemeine (meta)kognitive und affektive Variablen untersucht (enjoyment, cognitive load, judgement of learn-ing, involvement). Die gesamte Studie findet in einer Online-Umgebung statt.

Vorläufige Ergebnisse und ihre Bedeutung

Es liegen noch keine endgültigen Ergebnisse vor. Vorläufige Analysen geben Hin-weise darauf, dass Lernende im Hinblick auf Diagrammkompetenz nicht aber im Hinblick auf direkte Erinnerungsleistung vom schrittweisen Aufbau der Diagramme im Video profitieren. Weiterhin gibt es in der vorläufigen Analyse Indizien dafür, dass Lernende die Freude beim Lernen für das Video mit schrittweisem Video höher einschätzen.

Session G

Einzelbeiträge

G-01: Persönlichkeit und Kompetenz in verschiedenen Domänen

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S06 S00 B29
Chair der Sitzung: Patrick Fesser

Wissenschaftspropädeutik im Fach Mathematik: Erfassung der Fähigkeit zur meta-wissenschaftlichen Reflexion von Studienanfängerinnen und -anfängern

Patrick Fesser, Stefanie Rach
Otto-von-Guericke-Universität Magdeburg, Deutschland

Wissenschaftspropädeutik gehört neben vertiefter Allgemeinbildung und allgemeiner Studierfähigkeit zur Zieltrias der gymnasialen Oberstufe (KMK, 2021). Dabei kann Wissenschaftspropädeutik als eine Anbahnung einer wissenschaftlichen Denk- und Arbeitsweise verstanden werden. Zur Konkretisierung von Wissenschaftspropädeutik schlägt Müsche (2009) ein 3-Ebenen-Modell vor, welches hierarchisch aufgebaut ist: (1) meta-wissenschaftliches Wissen, (2) Methodenbewusstsein und (3) meta-wissenschaftliche Reflexion. Meta-wissenschaftliche Reflexion umfasst dabei die Reflexion von Wissenschaft aus einer epistemologischen Perspektive (Neubrand, 2000). Dazu gehört u. a. die Orientierung im System der wissenschaftlichen Disziplinen. Während bereits erste Ansätze zum Erfassen der meta-wissenschaftlichen Reflexion bzgl. des naturwissenschaftlichen Erkenntnisprozesses existieren (Kramer et al., 2021; Oschatz et al., 2018), ist die Literaturlage im Bereich Mathematik spärlich ausgeprägt. Ein Grund dafür liegt möglicherweise darin, dass dem Mathematikunterricht zwei wissenschaftspropädeutische Funktionen zugewiesen werden: Mathematik als deduzierend-beweisende und als anwendbare Disziplin erfahrbar zu machen (von Hentig, 1980).

In dieser Studie wird die Fähigkeit zur meta-wissenschaftlichen Reflexion von Studienanfänger*innen bezogen auf Mathematik als deduzierend-beweisende Disziplin untersucht. Die Forschungsfragen dieser Studie lauten:

(FF1) Wie wird Mathematik als wissenschaftliche Disziplin von den Studienanfänger*innen kategorisiert?

(FF2) Welche Zusammenhänge bestehen zwischen der Kategorisierung von Mathematik und kognitiven und affektiven Personenmerkmalen?

Eine Onlinebefragung mit 313 Studierenden aus mathematikhaltigen Studiengängen, z. B. BWL, Ingenieurpädagogik, wurde zu Beginn des Wintersemesters 2020/21 durchgeführt. Zur Erfassung der meta-wissenschaftlichen Reflexion sollten die Studierenden an die Mathematik an der Hochschule denken und Stellung zur Aussage „Mathematik ist eine Naturwissenschaft“ nehmen, Pro- und Kontra-Argumente sammeln und diese gegeneinander abwägen. Insgesamt haben 296 Studierende (56,1% weiblich, 82,4% jünger als 22 Jahre) eine Antwort auf das offene Item formuliert. Die Antworten wurden mithilfe eines entwickelten Kategoriensystems (Fesser & Rach, 2022) von zwei Ratern kodiert. Die Kategorisierung, ob die Studierenden Mathematik als Naturwissenschaft ansehen oder nicht, kann als reliabel eingeschätzt werden (korrigiertes Cohens κ_n = 0,93 nach Brennan & Prediger, 1981). Daneben wurden mithilfe eines Fragebogens kognitive (Abiturnote und letzte Mathematiknote im Punktesystem) und affektive Personenmerkmale (Interesse und Selbstkonzept bzgl. Beweisen)

abgefragt. Insgesamt liegen zu 155 Studierenden (56,1% weiblich, 86,5% jünger als 22 Jahre) Daten zu allen Merkmalen vor.

Zur Beantwortung von FF1 wurde inhaltsanalytisch untersucht, ob Studienanfänger*innen Mathematik als Naturwissenschaft kategorisieren bzw. welche Kategorien verwendet werden, um Mathematik als wissenschaftliche Disziplin zu charakterisieren. Von den 296 Studierenden kategorisieren 137 Studierende Mathematik als Naturwissenschaft, 140 Studierende Mathematik *nicht* als Naturwissenschaft und 19 Studierende haben sich nicht eindeutig positioniert. Von den 140 Studierenden, die Mathematik nicht als Naturwissenschaft bezeichnen, wurden die folgenden Beschreibungen gewählt: „Hilfswissenschaft" (24 Studierende), „besondere Wissenschaft" (10 Studierende), „Humanwissenschaft" (5 Studierende) und „Formalwissenschaft" (2 Studierende).

Zur Beantwortung von FF2 wurden die kompletten Fälle (N = 155) mithilfe einer multiplen logistischen Regression nach der Stepwise-Methode analysiert. Es zeigte sich, dass sowohl die Abiturnote (B = 0,66, SE = 0,32, p < 0,05) als auch die letzte Mathematiknote (B = 0,10, SE = 0,05, p < 0,05) signifikante Prädiktoren für die Kategorisierung von Mathematik als Naturwissenschaft sind. Betreffend der affektiven Merkmale zeigte sich nur das Interesse bzgl. Beweisen (B = -0,46, SE = 0,27, p = 0,09) als tendenziell signifikant. Die Prädiktoren besitzen eine geringe bis mittlere Erklärungskraft (Nagelkerke R^2 = 0,12).

Insgesamt scheinen sich die Studienanfänger*innen bzgl. der Kategorisierung von Mathematik als wissenschaftliche Disziplin uneinig zu sein. Dass Mathematik als deduzierend-beweisende Disziplin fälschlicherweise als Naturwissenschaft angesehen wird, ist bei der Hälfte der Studierenden der Fall. Entgegen unserer Erwartungshaltung geht eine gute letzte Mathematiknote damit einher, Mathematik eher als Naturwissenschaft zu kategorisieren. Dies könnte darauf hindeuten, dass der Mathematikunterricht Mathematik eher als anwendbare und nicht als deduzierend-beweisende Disziplin fokussiert. Durch die Argumente, die die Studierenden zur Entscheidung nutzen, kann die Fähigkeit zur Reflexion noch genauer gefasst werden und diskutiert werden, inwieweit Wissenschaftspropädeutik im Mathematikunterricht der Oberstufe eine realistische Zieldimension ist.

Disentangling the Association Between Big Five Personality Traits and Student Achievement: Meta-Analytic Evidence on the Role of Domain Specificity and Achievement Measures

Jennifer Meyer[1], Thorben Jansen[1], Nicolas Hübner[2], Oliver Lüdtke[1]
[1]IPN - Leibniz Institute for Science and Mathematics Education, Germany; [2]University of Tübingen

Students' academic achievement is a central predictor of a long list of important educational outcomes, such as access to higher education and socioeconomic success. Prior studies have extensively focused on identifying variables that are related to academic achievement and an important variable in this context appears to be students' personality (for prior meta-analyses see Poropat, 2009; Mammadov, 2022). Recent empirical findings (e.g., Brandt et al., 2021; Israel et al., 2019; Spengler et al., 2013) suggest that the association between student achievement and personality varies by subject domain (language vs. STEM) and the type of achievement measure (grades vs. test scores). Learning in school takes place in different subject domains, such as students' first and second languages, mathematics, and sciences, and domain-specific achievement shapes student career decisions even beyond subject-specific expectancy beliefs and interests (Guo et al.,

2017) and, thus, can influence the decision to pursue careers in different fields (e.g., STEM; Hübner et al., 2017; Jansen et al., 2015). Further, a long list of different achievement measures exists, which typically distinguishes between grades and standardized test achievement (Willingham et al., 2002). PASH (personality–achievement saturation hypothesis; Hübner et al., 2022) provides a comprehensive framework to describe differences between different achievement measures. In this framework, it is argued that personality manifests itself in a more pronounced way ("higher personality saturation") under certain circumstances that are reflected in different achievement measures. The moderating role of measures and domains on how personality relates to achievement has not yet been considered in previous meta-analyses.

Research Questions

We examined three main research questions. We investigated whether the subject domain (RQ1) and type of achievement measure (RQ2) moderates the association between personality traits and student achievement. In RQ3, we investigated whether there is evidence for a two-way moderation effect of academic subject domain and type of achievement measure (i.e., subject domain × achievement measure) on the effect sizes (i.e., the correlation between personality trait and achievement). In other words, we investigated whether the effects found between subject domains are similar across measures or whether the effects found between subject domains are enhanced or reduced in grades or standardized test scores, respectively.

Methods

To address this gap in the educational research literature, we conducted a meta-analysis based on 78 studies with 1,491 effect sizes representing data from 500,218 students and 110 samples from elementary to high school. We used a random-effects model with robust variance estimation to calculate mean effect sizes and standard deviations considering the nested data structure with studies providing multiple effect sizes. Literature search is illustrated in Figure 1.

Results

The results of our meta-analysis extend the findings of previous meta-analyses, providing a systematic approach to understanding the association of personality traits and achievement in school in more detail. We found moderating effects of measure or domain for all five personality traits, with differences in the direction of the effects. A graphical overview of results can be found in Figure 2. Our results highlight the importance of the domain and measure when examining how personality traits relate to academic achievement in school. The combination of subject domain and achievement was also found to be relevant for all traits except conscientiousness. These findings emphasize that subject domains and types of achievement measures should be explicitly considered when investigating the personality saturation of student achievement. Our findings show that the associations of personality traits with achievement cannot be generalized across measures and subject domains. Future researchers investigating the role of personality traits in association with learning outcomes should specifically consider which aspects of achievement are relevant to their investigation and choose the appropriate achievement measure.

Kreatives Potenzial und das Selbst: Eine Meta-Analyse zum Zusammenhang zwischen Divergentem Denken und kreativitätsbezogenen Selbsteinschätzungsmaßen

Sabine Patzl[1], Jennifer Diedrich[1], Jakob Pietschnig[2], Doris Lewalter[1]
[1]Zentrum für internationale Bildungsvergleichsstudien (ZIB), Technische Universität München, Germany; [2]Institut für Psychologie der Entwicklung und Bildung, Fakultät für Psychologie, Universität Wien, Österreich

Theoretischer Hintergrund:

Kreativität ist ein wesentlicher Faktor, der es Schüler*innen erlaubt, sich an die ständig verändernde Umwelt anzupassen (OECD, 2019). Darüber hinaus korreliert Kreativität positiv mit der akademischen Leistung (Gajda et al., 2017). Nicht nur Künstler*innen oder Profimusiker*innen verfügen über kreative Fähigkeiten, vielmehr ist es wissenschaftlicher Konsens, dass jede Person über ein gewisses Maß an kreativem Potenzial verfügt (Runco, 2003). Das kreative Potenzial einer Person wird häufig über das erstmals von Guilford (1950) beschriebene „Divergente Denken“ (= DT) erfasst. Neben Leistungsmaßen (z.B. der Torrance Tests of Creative Thinking zur Erfassung von DT; Torrance, 1966) wird Kreativität in der empirischen Forschung auch mittels Selbsteinschätzungsmaßen erhoben („creative self-assessments“ = CSA). Hierbei bewerten Proband*innen verschiedenste Facetten ihrer Kreativität selbst (für eine Überblick siehe Kaufman, 2019), wie etwa die Einschätzung der eigenen allgemeinen sowie domänenspezifischen kreativen Fähigkeit. Die Fragen können sich ebenfalls auf den kreativen Prozess, beispielweise bei der Ideengenerierung, oder auf das Auftreten bestimmter kreativer Aktivitäten, fokussieren. Die letzte Facette bezieht sich darauf, wie Personen Kreativität sehen, also welche Überzeugungen diese z.B. hinsichtlich der Veränderbarkeit von Kreativität in sich tragen. Jedoch gibt es Hinweise darauf, dass das kreative Potenzial einer Person nicht zwangsläufig auch zu kreativen Handlungen bzw. Errungenschaften führt, sondern, dass der Zusammenhang durch CSA, wie z.B. die Überzeugungen bezüglich Kreativität (Karwowski & Beghetto, 2019), beeinflusst wird.

Fragestellung:

Auf der Basis des geschilderten theoretischen Hintergrunds befasst sich diese Studie daher mit der Frage, inwieweit kreatives Potenzial (erfasst durch DT) in Zusammenhang mit CSA steht. Darüber hinaus wird untersucht, inwieweit dieser Zusammenhang z.B. durch das Alter der Personen (Vergleich von drei Altersgruppen: Kinder, Jugendliche und Erwachsene) oder die erfasste CSA Facette (bspw. Vergleich Prozess oder Überzeugung) moderiert wird.

Methode:

Der Zusammenhang zwischen CSA und Divergentem Denken wird mittels eines dreistufigen metaanalytischen Modells untersucht. Der Einfluss der erfassten CSA Facette sowie der Altersgruppe wird mittels Subgruppenanalysen analysiert. Schließlich werden mit Hilfe von Multiversums- und Spezifikationskurvenanalyse diverse meta-analytische Gesamteffekte verglichen, abhängig davon wie die Daten ("how factors") und welche Daten ("which factors") in den Primärstudien analysiert wurden (Voracek et al., 2019). Dies erlaubt uns die Stabilität bzw. die Robustheit des Zusammenhangs abhängig von unterschiedlichen aber gleichermaßen sinnvollen Spezifikationen zu schätzen.

Ergebnisse und ihre Bedeutung:

Die systematische Literaturrecherche in den Datenbanken PsychINFO, PSYNDEX, PsychARTICLES, Scopus, ISI Web of Science und ERIC ergab auf Basis der Keywords, Abstracts und Volltexte bisher 1714 potenziell relevante Studien. Diese werden derzeit gescreent und kodiert. Erste Ergebnisse werden im Rahmen der Konferenz berichtet werden können. Basierend auf theoretischen Überlegungen erwarten wir einen kleinen bis moderaten Zusammenhang zwischen DT und CSA, der abhängig von der untersuchten CSA Facette und der Altersgruppe variiert. Durch die Anwendung der Multiversums- und Spezifikationskurvenanalyse wird darüber hinaus eine Aussage darüber getroffen werden können, ob (und wie) sich der meta-analytische Gesamteffekt abhängig von unterschiedlichen Spezifikationen verändert und in welcher Stärke sich Effekte typischerweise beobachten lassen. Unsere Studie leistet somit einen wertvollen Beitrag zum Verständnis dessen, wie sich das kreative Selbstbild formt und in der Folge wie die Grundlage für kreative Leistungen in verschiedenen Altersstufen gelegt wird.

Entwicklung eines Kompetenzstrukturmodells zur Erfassung von Klimakompetenz am Ende der Sekundarstufe I

Magdalena Stadler, Monika Martin, Werner Rieß, Josef Künsting
Pädagogische Hochschule Freiburg, Deutschland

Das Voranschreiten des anthropogenen Klimawandels stellt die Welt vor beispiellose Herausforderungen, die nur mit gemeinsamer Anstrengung großer Teile der Weltbevölkerung bewältigt werden können. Dafür muss auch das Bildungssystem das Thema Klimawandel und die Bewusstseinsbildung in den Fokus nehmen. Klima und Klimawandel werden bisher allerdings zu wenig im deutschen Bildungssystem adressiert (Siegmund et al., 2021). Um diesem Defizit zu begegnen, sollte als Erstes der aktuelle Stand der Klimabildung an Schulen, hier insbesondere der Leistungsstand der Schüler*innen, ermittelt werden, um daraus adaptive und wirksame Maßnahmen zur Weiterentwicklung des Bildungssystems ableiten zu können (Drechsel et al., 2020).

Vor dem Hintergrund der dargestellten Situation stellt sich die übergeordnete Frage, wie Klimakompetenz bei Schüler*innen adäquat gemessen werden kann. Um die interessierende Kompetenz valide messen zu können, ist zunächst die Entwicklung eines Kompetenzmodells als theoretische Basis erforderlich (Klieme & Hartig, 2007, S. 24). Eine Herausforderung dabei ist, dass das Konstrukt der Klimakompetenz in der empirischen Bildungsforschung bisher wenig adressiert wurde und aus der Perspektive zahlreicher Fächer bestimmt werden muss. Dieser Forschungslücke widmet sich der geplante Beitrag:

Die interdisziplinäre Forschungsgruppe der Autor*innen entwickelt einen psychometrischen Test zur Erfassung von Klimakompetenz bei Schüler*innen am Ende der Sekundarstufe I (9. Klasse). Grundlage hierfür bildet ein Kompetenzstrukturmodell, das die kognitiven Komponenten von Klimakompetenz in den relevanten Facetten abbildet. Bei der theoretischen Herleitung des Kompetenzmodells greifen wir auf Theorien, Modelle und Konstrukte aus Umweltpsychologie (z.B. Roczen et al., 2014) und empirischer Bildungsforschung zur BNE (z.B. Waltner et al., 2019; Wiek et al., 2011) zurück. Außerdem orientieren wir uns an den Kompetenzmodellen, die in den großen Schulleistungsstudien PISA, TIMMS und IQB für die naturwissenschaftliche Domäne konzeptualisiert wurden (z.B. OECD, 2017).

Klieme und Leutner (2006, S. 879) folgend definieren wir „Kompetenzen als kontextspezifische kognitive Leistungsdispositionen, die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen." Im Sinne des Kompetenzbegriffs nach Weinert (2001) gehen wir zudem davon aus, dass motivationale Orientierungen, Einstellungen, Tendenzen und Erwartungen auch für den Umgang mit dem Klimawandel von hoher Bedeutung sind. Daraus leiten wir das Konstrukt der Climate Literacy (Klimagrundbildung) ab, die nach unserem Verständnis Klimakompetenz in Form von Kenntnissen und Fähigkeiten zum Klimawandel in Kombination mit motivationalen und volitionalen Orientierungen umfasst, um Problemlösungen und Handlungen mit Relevanz für den Klimawandel zu ermöglichen. Die motivational- affektiven Komponenten werden neben dem Wissen und den kognitiven Fähigkeiten separat in zusätzlichen Skalen erhoben, um eine präzise Operationalisierung des Kompetenzstrukturmodells einhalten zu können (Klieme & Hartig, 2007, S. 25).

Das Kompetenzstrukturmodell beinhaltet zwei Dimensionen:

1. Dimension „Kompetenzbereiche": Aus unserer Definition von Klimakompetenz lassen sich vier Kompetenzbereiche ableiten, die als grundlegend für einen adäquaten Umgang mit dem Klimawandel angesehen werden können:

1. (1) Umgang mit Fachwissen
2. (2) Gewinnung und Beurteilung von Erkenntnissen
3. (3) Information und Kommunikation
4. (4) Normative Bewertung

2. Dimension „Inhalte": Die Konzipierung fundierter theoretischer Kompetenzmodelle als Grundlage für empirische Messverfahren „erfordert die gemeinsame Anwendung von pädagogisch-psychologischer und fachdidaktischer Expertise in der jeweiligen Domäne" (Klieme & Hartig, 2007, S. 26). Diesem Anspruch folgend, haben neun Fachdidaktiker*innen verschiedener Disziplinen zwölf Basiskonzepte der Climate Literacy (z.B. zum „Klimasystem und naturwissenschaftlichen Grundlagen") entwickelt. Diese wurden von zehn Expert*innen (z.B. Fachwissenschaftler*innen aus Klimatologie, Geografie, Politik) validiert.

Neben den zwei vorgestellten Kompetenzdimensionen werden bei der Operationalisierung unter anderem fünf kognitive Prozesse nach der Revision der Bloom'schen Lernzieltaxonomie nach Krathwohl (2002) und die Einbettung der Testaufgaben in verschiedene Kontexte berücksichtigt, um eine Ausgewogenheit in der Schwierigkeit zu gewährleisten.

Wesentliches Ziel des geplanten Beitrags für die GEBF-Tagung ist es, die wissenschaftlich fundierte Entwicklung eines Kompetenzstrukturmodells vorzustellen, das als theoretische Grundlage zur Operationalisierung eines psychometrischen Tests zur Erfassung von Klimakompetenz dient. Außerdem sollen Daten zur ersten Pilotierung des Testinstruments präsentiert werden, die Ende 2022 erfolgt.

G-02: Beziehungen und Strukturen

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S06 S00 B32
Chair der Sitzung: Julia Künstler-Sment

Fördern Ganztagsangebote in der Sekundarstufe I diverse Freundschaftsbeziehungen und Outgroup-Kontakte?

Julia Künstler-Sment, Aleksander Kocaj, Georg Lorenz, Malte Jansen
Institut zur Qualitätsentwicklung im Bildungswesen, Deutschland

Während der Adoleszenz werden Peers für die Jugendlichen immer wichtiger, und sie verbringen auch mehr Zeit miteinander im Unterricht und bei außerunterrichtlichen Aktivitäten in schulischen Kontext (Eccles et al., 2003). Ein Ziel des Ausbaus von Ganztagsangeboten liegt darin, positive und diverse soziale Beziehungen zwischen Schüler:innen zu fördern. Die Entwicklung sozialer Beziehungen zwischen Jugendlichen hängt auch davon ab, wie ähnlich sich die Personen untereinander in Hinblick auf bestimmte Hintergrundmerkmale wie Geschlecht, aber auch ethnischer und sozialer Hintergrund sind (Lorenz et al., 2020).

Die Personen, die sich in bestimmten Merkmalen (wie z.B. Ethnizität) ähnlich sind, bilden die Eigengruppe, während die Personen, die sich in diesen Merkmalen unterscheiden, eine Fremdgruppe formen (Billig & Tajfel, 1973). Pflegt eine Person die Beziehung zu einer anderen Person aus der Eigengruppe, handelt es sich um Ingroup-Kontakt, die Beziehung zu einer Person aus der Fremdgruppe hingegen wird als Outgroup-Kontakt definiert (Wölfer et al., 2019).

Soziale Netzwerke der Jugendlichen mit Migrationshintergrund bieten nur begrenzte Möglichkeiten für den Zugang zu Sozialkapital (Stanat et al., 2010). Das Sozialkapital stellt die Ressourcen dar, die über die Beziehungen innerhalb des sozialen Netzwerks verteilt sind (Bourdieu, 1986) und zu denen eine Person durch direkte oder indirekte soziale Kontakte Zugang hat (Lin, 2001). Je mehr verschiedene Kontakte man unterhält, desto wahrscheinlicher ist es, dass einer von ihnen über die benötigte Ressource verfügt (Erickson, 2003).

Im schulischen Umfeld können extracurricularen Aktivitäten einen besonders förderlichen Rahmen für positive Kontakte zwischen verschieden Gruppen bieten (Moody, 2001). Nach Allport (Allport, 1985) sollten in sozialen Kontaktsituationen positive soziale Interaktionen (z. B. Freundschaftsentwicklung) zwischen Gruppen zustande kommen, wenn (1) die Gruppen in der Kontaktsituation gleichgestellt sind, (2) gemeinsame Ziele angestrebt werden, (3) eine Kooperation zwischen den Gruppen stattfindet und (4) der Kontakt durch öffentliche Institutionenunterstützt wird – alle diese Kriterien werden im Rahmen von extracurricularen Aktivitäten erfüllt. Extracurriculare Aktivitäten finden in Deutschland verstärkt an Ganztagsschulen statt, die auch im Schnitt eine diverse Schülerschaft aufweisen (Berkemeyer, 2015).

Die zentrale Hypothese lautet, dass Jugendliche an Ganztagsschulen, die an diversen Ganztagsangeboten teilnehmen und häufiger mit Gleichaltrigen interagieren können, dadurch mehr Freundschaften über Kategoriengrenzen wie z. B. Herkunftsland entwickeln als Jugendliche an Halbtagsschulen. Die Analysen basieren auf den Daten des IQB-Bildungstrends 2018 (Stanat et al., 2022). 11.726 Schüler:innen (15,49 Jahre alt; 49,6% Mädchen) aus 549 Klassen in 427 deutschen Sekundarschulen (261 Halbtag- und 166 Ganztagschulen) gaben Auskunft über ihre sozialen Beziehungen im Klassenverbund. Die Freundschaftsnetzwerke innerhalb der Schulklassen wurden mit einer soziometrischen Nominierungsmethode auf Basis einer vollen Klassenliste erfasst. Die

zentrale abhängige Variable wurde gebildet, indem die Anzahl der Outgroup-Kontakte (Beziehungen zu den in einem anderem Land geborenen Mitschüler:innen) von der Anzahl der Ingroup-Kontakte (Beziehungen zu den im gleichen Land geborenen Mitschüler:innen) subtrahiert und durch die Gesamtzahl der Kontakte pro Person geteilt wurde (Wölfer et al., 2019). Um Unterschiede auf Individualeben und in der Zusammensetzung der Schüler*innenschaft an Halbtagsschulen und an Ganztagsschulen zu kontrollieren, welche unterschiedliche Chancen für Outgroup-Kontakte bereitstellen, wird ein Propensity Score Matching-Verfahren (Full Matching) eingesetzt.
Die Ergebnisse weisen darauf hin, dass in Halbtags- und Ganztagsschulen die in Deutschland geborene Schüler:innengruppe im Durchschnitt mehr Ingroup-Kontakte hatte, während die nicht in Deutschland geborene Schüler:innengruppe mehr Outgroup-Kontakte aufwies. Die Effektstärken deuten darauf hin, dass nicht in Deutschland geborene Schüler:innen in Ganztagsschulen weniger Outgroup-Kontakte hatten als in Halbtagsschulen, während in Deutschland geborene Schüler:innen in Ganztagsschulen weniger Ingroup-Kontakte hatten als in Halbtagsschulen. Nach dem Full Matching zeigte sich ein ähnliches Befundmuster, welches nicht hypothesenkonform war: ein positiver Zusammenhang zwischen Ganztagsschulbeschulung und der Tendenz zu Ingroup-Kontakten. Differenziert nach außerunterrichtlichen Angeboten zeigen sich sowohl positive als auch negative Effekte auf die Tendenz zu Outgroup-Kontakten, die in weiterführenden Analysen genauer untersucht (z. B. Einfluss der Schüler:innenleistung, Kontaktmöglichkeiten in Angeboten) werden.

Profile von Geschwisterbeziehungen und ihr Zusammenhang mit sozial-emotionalen Fähigkeiten

Elena Wittmann, Ilka Wolter
Leibniz-Institut für Bildungsverläufe, Deutschland

Die Geschwisterbeziehung ist eine der am meisten vernachlässigten Beziehungen in der bildungswissenschaftlichen Forschung. Das ist erstaunlich, denn aufgrund der Häufigkeit und des Umfangs der Interaktionen, der Dauerhaftigkeit der Beziehung, des Vorhandenseins zugeschriebener Rollen und des Ausmaßes an gemeinsamen Erfahrungen bilden diese Beziehungen einen wichtigen Kontext, in dem Kinder sozial-emotionale Fähigkeiten entwickeln.

Trotz des Fehlens einer umfassenden Geschwistertheorie tauchen in der Literatur immer wieder mehrere Dimensionen zur Beschreibung von Geschwisterbeziehungen auf, wie Wärme, Feindschaft oder Rivalität. Die Zusammenhänge zwischen diesen Dimensionen der Qualität der Geschwisterbeziehung und sozial-emotionalen Fähigkeiten können durch verschiedene Theorien erklärt werden: Ein Ansatzpunkt sind die theoretischen Grundlagen der Bindungstheorie (Ainsworth, 1989). Stabile Bindungsbeziehungen zu Geschwistern können das Sicherheitsgefühl des Kindes stärken und als Schutzfaktor wirken. Eine negative Bindungserfahrung hingegen kann zu niedrigem Selbstwertgefühl und Problemverhalten führen (Buist et al., 2013). Eine zweite Theorie, welche die Auswirkungen von Geschwistern auf sozial-emotionale Fähigkeiten erklären kann, ist die Theorie des sozialen Lernens (Bandura, 1977). Diese stellt dar, wie Interaktionsmuster in der Geschwisterbeziehung auf andere Kontexte übertragen werden können. Ein dritter theoretischer Zugang ergibt sich durch die Theorie des sozialen Vergleichs (Festinger, 1954) die beschreibt, dass Personen dazu neigen, sich mit anderen zu vergleichen und ihr Selbstwertgefühl aus diesem

Vergleich abzuleiten. Angesichts ihres gemeinsamen familiären Hintergrunds und ihrer gemeinsamen Erfahrungen sind Geschwister eine unmittelbare Referenzgruppe für soziale Vergleiche.

Gleichzeitig sind Geschwisterbeziehungen emotionsgeladene Beziehungen, die durch starke, ungehemmte Emotionen von positiver, negativer und manchmal auch ambivalenter Qualität gekennzeichnet sind. Auf der Grundlage der beiden Dimensionen Wärme und Konflikt wurden in bisheriger Forschung vier Arten von Geschwisterbeziehungen gefunden: harmonische, feindselige, affektintensive und distanzierte (z.B. Derkman, 2011). Angesichts der Häufigkeit von Konflikten zwischen Geschwistern könnte es wichtig sein, die Art und Weise zu untersuchen, wie Konflikte in dieser Beziehung gelöst werden. Daher ist eine zentrale Forschungsfrage dieses Beitrages, ob die Arten der Geschwisterbeziehung unter Berücksichtigung weiterer Dimensionen (Rivalität und Konfliktresolution) repliziert werden können. Des Weiteren sollen auf Grundlage der gefunden Arten die Zusammenhänge mit sozial-emotionale Fähigkeiten herausgearbeitet werden.

Zu diesem Zweck werden Daten des Beziehungs- und Familienpanels *pairfam* (Brüderl et al., 2021) genutzt. Insgesamt konnten 652 Kinder (50,77 % männlich) mit einem Altersdurchschnitt von etwa 11 Jahren (*SD*=2,1) identifiziert werden, welche mindestens ein Geschwisterkind hatten und zu dieser Beziehung Angaben getätigt haben. In einem ersten Schritt wurden latente Profilanalysen durchgeführt. Hierfür wurden drei Skalen des Sibling Relationship Inventorys (Stocker & MacHale, 1992; Boer et al., 1997) und eine Skala zur Konfliktresolution der Sibling Qualities Scale (Cole & Kerns, 2001) verwendet. Als Indikatoren für sozial-emotionale Fähigkeiten dienen die Subskalen prosoziales Verhalten und Verhaltensprobleme des Strengths and Difficulties Questionnaires (Goodman et al, 1998) und eine Skala zum Selbstwertgefühl (Rosenberg, 1965).

Die vier bisher gefundenen Arten der Geschwisterbeziehung konnten bestätigt und um eine weitere Art, die sich insbesondere durch hohe Rivalität auszeichnet, ergänzt werden. Im nächsten Schritt wurde die Profilzugehörigkeit durch multinomiale logistische Regression überprüft, dabei zeigte sich u.a., dass Schwesternpaare häufiger in harmonischen Geschwisterbeziehungen zu finden sind. Im letzten Schritt wurde eine Regression mit der Profilzugehörigkeit als Prädiktor für sozial-emotionale Fähigkeiten gerechnet. Kinder mit einer harmonischen Geschwisterbeziehung berichteten über ein höheres Maß an prosozialem Verhalten, ein geringeres Maß an Verhaltensproblemen und ein höheres Selbstwertgefühl. Kinder mit einer konfliktreichen Geschwisterbeziehung berichteten über ein geringeres Maß an prosozialem Verhalten, ein höheres Maß an Verhaltensproblemen und ein geringeres Selbstwertgefühl. Kinder mit einer rivalisierenden Geschwisterbeziehung berichteten über ein höheres Maß an Verhaltensproblemen und ein geringeres Selbstwertgefühl. Kinder mit einer affektintensiven Geschwisterbeziehung berichteten über ein geringeres Selbstwertgefühl. Hier scheint sich der Konflikt unabhängig vom Ausmaß der Zuneigung nachteilig auszuwirken. Ergänzend zu den Belegen um den Einfluss von Eltern-Kind-Beziehungen auf die Entwicklung von Kindern beleuchtet die Diskussion dieses Beitrags die Bedeutung der Geschwisterbeziehung.

Understanding the Composition of the Student Body - Cluster Analysis of Students in Higher Education Institutions

Thomas Gößl, Leandro Henao, Luis Rumert
University of Wuppertal, Germany

Results from educational data mining can be used to improve educational outcomes. For instance, machine learning algorithms have made it possible to design early detection systems that identify students at risk of dropping out, before the availability of first performance data (Berens et al., 2019; Del Bonifro et al., 2020; Lykourentzou et al., 2009). As the student population becomes more heterogeneous, the focus should not only be on predicting study success or dropout but also on identifying factors explaining these outcomes. Using standardised individual-level data from a state university in North-Rhine Westphalia, we implement Uniform Manifold Approximation and Projection (UMAP), a state-of-the-art non-linear dimensionality reduction technique, together with Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), an unsupervised machine-learning algorithm to cluster the student population. This procedure supports the identification of demographic, academic, and performance information that contribute the most to student attrition and allows for a cluster-specific interventions' implementation.

There are numerous situations where it is useful to classify or segment a given population. Customer segmentation using machine learning is one of the most common examples, where companies aim to address more homogenous groups of clients and optimise churning rates (Dullaghan & Rozaki, 2017; Smeureanu et al., 2013). Analogously, in the field of learning analytics and educational data mining, researchers are interested in predicting academic trends and patterns to sketch students (Parack et al., 2012), classifying e-learning behaviour (Chen et al., 2009) or identifying variables influencing performance and prescribing interventions (Salazar et al., 2004). However, most of these studies exhibit significant data limitations and are restricted to classical *k*-means and fuzzy clustering algorithms.

Naturally, having many variables challenges the clustering process and is computationally demanding. Therefore, we first apply the UMAP algorithm (McInnes et al., 2018). In the next step, we apply density-based clustering. In contrast to other clustering algorithms, HDBSCAN (Campello et al., 2013) is less sensitive to noisy data. Therefore, the procedure can be applied to analyse the composition of the student body independent of the institutional context and help to better understand student behaviour in various institutions and target planned interventions at the institutions of higher education. Finally, the combination of HDBSCAN with UMAP improves the clustering performance accuracy and speed (Allaoui et al., 2020).

Identified clusters do not only differ by demographic information but also according to academic achievements and activity level of the students. We find the expected student clusters, for example: male dropouts with and without migration background or female graduates born in Germany. Additionally, we identify clusters that have not been focussed on so far. Examples are: dropouts who are active students at the beginning of their studies but turned inactive later on, or graduates with low activity and low grades. Based on student data, our method differentiates various clusters of dropouts as well as clusters of graduates versus dropouts. Moreover, we identify clusters of currently active students that are more similar to either graduate or dropout clusters. This measure of closeness can be used to indicate a possible change of status (from student to graduate or dropout). This insight can be helpful for targeting support for the students. Moreover, to make the analysis more useful for

practitioners, we categorize and label the larger clusters to provide a more general framework of the student body composition.

Tuition Fees and Academic (In)Activity in Higher Education. How Did Students Adjust to the Abolition of Tuition Fees in Germany?

Johannes Berens[1], Leandro Henao[1], Kerstin Schneider[2]
[1]Bergische Universität Wuppertal, Germany; [2]Bergische Universität Wuppertal, Germany, CESifo

Only five years after introducing student fees, the German state of North-Rhine Westphalia (NRW) issued its abolition in March 2011. Using student-level panel data, we assess the effects on observed academic activity in two universities in NRW: a state university (SU) and a private university of applied sciences (PUAS). We estimate the impact of this reform on different student outcomes, employing a novel estimator that addresses important limitations of the traditional DID approach. Our analyses provide evidence that the abolition of tuition fees leads to decreased academic activity and performance among students at both institutions.

Bahrs & Siedler, 2019; Brahm et al., 2017; and Hübner, 2012 provide evidence that student fees might keep students from demanding higher education and, consequently, have adverse effects on social mobility and ethnic diversity (Allen & Wolniak, 2019). On the other hand, tuition fees might have an efficiency-enhancing effect by signalling the value of education (Beneito et al., 2018; Garibaldi et al., 2012; Sá, 2019). Furthermore, the financial benefits of the student status in Germany yield incentives to enrol at a university, regardless of the intention to complete the studies. These monetary incentives, such as a discounted public transportation ticket, distort public funds' allocation and the perception of successful and not successful student careers. If students enrol but never intend to complete the degree, they are eventually dropouts, even though they should be distinguished from students who aim at a degree but are unsuccessful.

This paper exploits the variation from abolishing moderate student fees of no more than EUR 500 per semester in a federal state of Germany. We use the administrative student panel data of two higher education institutions (HEI): a rich set of demographic and academic information on students. Most importantly, we have data on their performance and observed academic activity throughout the years they remain enrolled. Methodologically, we first cluster former students of both HEI according to their activity level, which we observe every time they pass, fail, withdraw, or never show up for an exam. Second, we employ a novel two-way fixed-effects (TWFE) estimator to investigate the effects of this reform in a staggered treatment introduction framework. We address treatment effect heterogeneity by utilising the estimator proposed by Callaway & Sant'Anna (2021), which produces cohort-time-specific treatment effects through simple 2x2s DID estimators. This empirical approach allows the estimation of ATTs under more general conditions than in the classical TWFE methods, such as when the parallel-trends assumption holds only conditionally.

We contribute to the existing literature by providing new insights into the concept of study and dropout motivation (Heublein et al., 2020) and analysing inactivity in the German higher education system (Bietenbeck et al., 2021). Most studies regarding student fees in Germany focus on their introduction before the Bologna process had been completed and utilise aggregated data (Bruckmeier & Wigger, 2014; Mitze et al., 2015; Thomsen & von Haaren-Giebel, 2016). Therefore, our use of student-level

panel data in the Bachelor system is a remarkable contribution. The abolition of tuition fees is linked to decreased academic activity among students. The effects of such a policy substantially differ not only concerning the type of institution -private or public- where it is implemented, but also with respect to demographic dimensions. Eliminating tuition fees correlates with a sizeable increase in total inactivity or ghost students (40%). Excluding cohort composition changes, we find that students complete up to 25% less CP per semester than tuition-paying individuals. This is driven by reduced academic effort in registered and passed exams.

G-03: Mathematikunterricht

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S06 S00 B41
Chair der Sitzung: Nadine Besser

Frühes Zahlenwissen und die längsschnittlichen Beziehungen zu (mathematischen) Sprachkenntnissen

Nadine Besser[1], Simone Lehrl[2], Anja Linberg[3], Dorothea Dornheim[1], Sabine Weinert[1]
[1]Otto-Friedrich-Universität Bamberg, Deutschland; [2]Pädagogische Hochschule Weingarten, Deutschland; [3]Deutsches Jungendinstitut, Deutschland

Bereits beim Eintritt in die Kindertageseinrichtung (Kita) gibt es erhebliche interindividuelle Unterschiede in frühen numerischen Kompetenzen von Kindern, die den Grundstein für spätere Kompetenzen legen (Lehrl et al., 2020). Neben der Kita-Qualität (Lehrl et al., 2016) und verschiedenen Hintergrundfaktoren (Jordan & Levine, 2009) gilt die sprachliche Entwicklung als besonderer Katalysator für die numerische Entwicklung (Sarnecka, 2014). Bisherige querschnittliche Befunde deuten auf eine positive Beziehung zwischen der mathematischen Sprache und den numerischen Kompetenzen von zweijährigen, auch über die allgemeinen Sprachkenntnisse hinaus (Autor*in et al., 2021). Die Richtung der Beziehung sowie die Rolle von familiären und institutionellen Einflussgrößen in diesem Gefüge ist dabei unklar.

Der vorliegende Beitrag untersucht daher die längsschnittlichen Zusammenhänge zwischen der mathematischen Sprache und den numerischen Kompetenzen der Kinder unter Berücksichtigung verschiedener Hintergrundfaktoren sowie der Interaktionsqualität, gemessen mit dem Classroom Assessment Scoring System (CLASS-Toddler; La Paro et al., 2012).

Methode:

Grundlage sind Daten der Studie EarlyMath, in der familiäre und institutionelle Faktoren, welche die mathematische Entwicklung von zwei- bis vierjährigen Kindern beeinflussen, im Längsschnitt untersucht werden. Das Projekt begann Anfang 2020 und ist als Interventionsstudie mit einem Prä-, Post, Follow-Up-Design konzipiert. In den beiden Experimentalgruppen erhielten die frühpädagogischen Fachkräfte eine mehrtägige Schulung zum Thema globale (EG1) bzw. mathematische (EG2) Interaktionsqualität; die Fachkräfte der beiden Kontrollgruppen (KG1 & KG2) lediglich einen Informationsnachmittag zu Interaktionsqualität. Alle Gruppen mit Ausnahme von KG2 erhielten zudem eine Spielzeugbox.

Die Stichprobe umfasst N=172 Kinder im Alter von durchschnittlich 32 Monaten (range = 23-41 Monate) zum ersten Messzeitpunkt (53 % männlich). Zweiundsiebzig Prozent der Kinder sprechen zu Hause ausschließlich Deutsch; etwas mehr als die Hälfte der Eltern haben eine akademische Ausbildung (57 %).

Messungen:

Die Zählfertigkeiten der Kinder wurden spielerisch mit den Aufgaben *Verbales Zählen* sowie *Objekte Abzählen* erfasst. Mit den Aufgaben *Point-to-X* (Wynn, 1992) und *Give-a-Number* (Sarnecka & Carey, 2008; Wynn, 1990) wurde gemessen, inwieweit das Zahlenkonzept bei den Kindern bereits entwickelt ist. Die Maße wurden zu einem Gesamtwert integriert (M=17,59; SD=7,36). Der mathematische Wortschatz der Kinder wurde mit einer erweiterten deutschen Version des Preschool Assessment of the Language of Mathematics (PALM; Purpura & Logan, 2015) getestet und beinhaltet 21 Items zum

Verstehen quantitativer und räumlicher Begriffe (M=8,79; SD=3,24). Der allgemeine Wortschatz wurde mit dem Peabody Picture Vocabulary Test (PPVT-IV; Dunn & Dunn, 2007) gemessen (M=37; SD=24,8). Um gezielt die Entwicklung der Kinder und die Anregungsqualität in der Kita zu adressieren, wurden zudem die numerischen Kompetenzen der Kinder sowie die Interaktionsqualität (CLASS-Toddler; La Paro et al., 2012) zum ersten Messzeitpunkt in die Analyse einbezogen.

Die Variablen Alter und Geschlecht der Kinder sowie Indikatoren des familiären Hintergrunds (höchster Bildungsgrad der Eltern, mittleres Familieneinkommen, Familiensprache) und die Region der Kita gehen als Kontrollvariablen in die Analysen ein.

Um der Frage nachzugehen, ob die mathematische Sprache der Kinder unter Berücksichtigung ihrer numerischen Kompetenzen zum ersten Messzeitpunkt, verschiedener familiärer Hintergrundfaktoren sowie der Einrichtungsqualität, längsschnittlich mit ihren späteren numerischen Kompetenzen zusammenhängen, werden schrittweise multiple Regressionen berechnet.

Ergebnisse:

Erwartungsgemäß sind das Alter (*ß*=.16, *p*<.050) sowie die numerischen Kompetenzen der Kinder im Pretest (*ß*=.41, *p*<.000) Prädiktoren für ihre numerischen Kompetenzen vier Monate später. Zudem erreichen Mädchen höhere Werte in den Aufgaben, als Jungen (*ß*=.27, *p*<.050), ebenso wie Kinder aus Familien mit einem höheren Einkommen (*ß*=.14, p<.050). Die Einrichtungsqualität (Emotional and Behavioral Support: *p*=.759, Engaged Support for Learning: *p*=.861), die Familiensprache (*p*=.616) sowie der rezeptive Wortschatz der Kinder (*p*=.117), gemessen im Pretest, erweisen sich als nicht prädiktiv.

Als Hauptbefund zeigt sich, dass die mathematische Sprache der Kinder – auch unter Kontrolle von Hintergrundvariablen, des rezeptiven Wortschatzes und der Interaktionsqualität – die Entwicklung ihrer numerischen Kompetenzen signifikant vorhersagt (*ß*=.19, *p*<.050). Das finale Modell erklärt 32,87 % der Unterschiede in den numerischen Kompetenzen der Kinder.

Formatives Assessment am Beispiel des Mathematikunterrichts: Potenziale und Herausforderungen des SMART-Online-Diagnosetools

Bärbel Barzel, Katrin Klingbeil, Fabian Rösken, Florian Schacht
Universität Duisburg-Essen, Germany

Formatives Assessment gilt als eines der wirksamsten Rahmenkonzepte zur Förderung schulischen Lernens (Schütze et al., 2018; Black & William, 1998; Kingston & Nash, 2011), es kommt aber z. B. aufgrund des erforderlichen Zeitaufwands (Bürgermeister & Saalbach, 2018) oder geringer Expertise und Erfahrung der Lehrkräfte (Abs, 2005; Bürgermeister, 2014) nicht im gewünschten Umfang zur Anwendung. Digitale Technologien können Lehrkräfte bei der anspruchsvollen und zeitaufwändigen Durchführung Formativer Assessments unterstützen. Um entsprechende diagnose- und förderorientierte Informationen zu liefern, reicht es jedoch nicht aus, wenn digitale Technologien nur die Richtigkeit der Antworten wiedergeben. Vielmehr sollten sie so konzipiert sein, das Denken der Lernenden zuverlässig zu erfassen.

Die forschungsbasierte Entwicklung der SMART-Tests (*Specific Mathematics Assessments that Reveal Thinking*) fußt auf entsprechenden Arbeiten von Stacey et al. (2018). Auf der Grundlage einer

automatischen Analyse von Mustern zwischen einzelnen diagnostischen Items liefert SMART eine Diagnose für jede:n einzelne:n Schüler:in in Form von Verstehensstufen und Fehlvorstellungen bzw. typischen Fehlern sowie entsprechend darauf abgestimmte kompetenzorientierte Förderhinweise (Steinle et al., 2009). Derzeit werden die SMART-Tests für den Einsatz im deutschsprachigen Raum adaptiert und implementiert, begleitet von einem groß angelegten Forschungsprojekt, das die Auswirkungen des Einsatzes von SMART-Tests exemplarisch zum Thema Variablenverständnis untersucht (Klingbeil et al., 2022).

Da Unterschiede in der Unterrichtskultur und in der Unterrichtspraxis die Vorstellungen der Lernenden beeinflussen können, ist es sinnvoll, die Eignung der in Australien entwickelten SMART-Tests für den Einsatz in Deutschland zu überprüfen. Der Beitrag adressiert daher die folgende Forschungsfrage: Inwieweit erfassen die ins Deutsche übersetzten Multiple Choice-Items der SMART-Tests zum Variablenverständnis die (Fehl-)Vorstellungen deutscher Schüler:innen angemessen?

Um sich dieser Frage zu nähern, wurden in einem ersten Schritt die entsprechenden SMART-Tests übersetzt und an deutsche Kontexte angepasst. Die Durchführung des Online-Tests erfolgte in einer 8. Klasse (n=19). Daraufhin wurden 11 halbstrukturierte Interviews mit ausgewählten Lernenden durchgeführt, bei denen parallele Aufgaben gelöst werden mussten. Ausgewählte Aufgaben wurden im Sinne offener Aufgaben umgewandelt, um mögliche Auswirkungen durch vorgegebene Antwortmöglichkeiten zu untersuchen. Anschließend wurden die Interviews qualitativ im Hinblick auf mögliche zugrunde liegende (Fehl-)Vorstellungen der Lernenden ausgewertet und mit der automatischen Online-Diagnose verglichen.

Im Allgemeinen wurde eine hohe Übereinstimmung zwischen den Ergebnissen des Online-SMART-Tests und den Erklärungen der Lernenden festgestellt. Abweichungen konnten vor allem bei der Fehlvorstellung "Lösung als Koeffizient" festgestellt werden (d.h. die Verwendung einer möglichen Lösung als Koeffizient in einer Gleichung, die das Problem beschreiben sollte): Sie wurde nicht in allen Fällen durch Multiple-Choice-Aufgaben aufgedeckt, sondern vereinzelt nur durch das Interview. Außerdem wurden unbeabsichtigte Schwierigkeiten in der übersetzten Version festgestellt, die sich aus der unterschiedlichen Verwendung bestimmter Buchstaben als Einheitssymbole ergeben. Auch zusätzliche Probleme, die in den Online-Tests nicht identifiziert wurden, wie z. B. implizite Multiplikationen, konnten im Rahmen der Interviews rekonstruiert werden.

Im Beitrag wird diskutiert, welche (Fehl-)Vorstellungen durch die Multiple-Choice-Aufgaben im Vergleich zu den Interviews aufgedeckt werden konnten, und es wird erörtert, wie die Aufgaben für das digitale Assessment-Tool angepasst werden können, um das Denken der Schüler:innen in einem groß angelegten Rahmen angemessen zu erfassen.

Brüche als Anteile verstehen mit digitalen Modellierungswerkzeugen

Rowena Merkel, Katharina Loibl, Timo Leuders, Frank Reinhold
Pädagogische Hochschule Freiburg, Germany

Theoretischer Hintergrund

Lernende bilden häufig nur ein unzureichendes konzeptuelles Wissen zentraler mathematischer Konzepte aus, da die Lerngelegenheiten stärker das prozedurale Wissen fokussieren (für Brüche, vgl.

Lenz et al., 2019; Reinhold, 2019). Im Kontext wirksamer Lerngelegenheit lassen empirische Studien den Schluss zu, dass digitale Lernumgebungen positiv den Lernerfolg beeinflussen können (Hillmayr et al., 2020), und besonders im Inhaltsbereich der Bruchrechnung und wenn dynamische Repräsentationen verwendet werden (Kiili et al., 2018; Reinhold et al., 2020). Dabei spielen die kognitiven Aktivitäten, die bei den Lernenden durch digital angereicherte Lerngelegenheiten ausgelöst werden, eine entscheidende Rolle (Rolfes, 2018; Roth, 2005, 2008).

Die Untersuchung solcher kognitiven Aktivitäten im Lernprozess ist methodisch herausfordernd. Daher wurde eine digitale Lernumgebung zur eigenständigen Erarbeitung des konzeptuellen Verständnisses des Größenvergleichs von Brüchen entwickelt. Die verschiedenen Bruchvergleichsaufgaben, bei denen durch technische Einschränkungen und/oder das verwendete Zahlenmaterial, jeweils unterschiedliche Strategien des Größenvergleichs von Brüchen getriggert werden, erlauben die Generierung einer Vielzahl an Lösungen und Strategien (Kapur & Bielaczyc, 2012; Loibl et al., 2017) als Indikator für die kognitive Aktivierung. Dies kann durch die Darbietung mehrerer Aufgaben (Kapur, 2010) oder durch externe dynamische visuelle Repräsentationen, die verschiedene Strategien zulassen, erreicht werden und ermöglicht eine nicht-reaktive Messung von konkreten kognitiven Aktivitäten während des Lernprozesses.

Die Lernumgebung liegt in zwei Varianten (dynamische vs. statische visuelle Repräsentation) vor, die sich darin unterscheiden, ob das Problem der Aufgabenstellung zunächst mit dynamischen Bruchstreifen (Prediger, 2013) selbstständig in ein Situationsbild überführt werden muss (dynamisch) oder, ob dieses Bruchstreifenbild bereits gegeben ist (statisch). Auf der Grundlage der visuellen Repräsentation der Bruchstreifen können Lernende eine Lösung für den Bruchvergleich finden sowie ein erstes Verständnis der Äquivalenz von Brüchen entwickeln. Erwartet wird, dass die selbstständige Erarbeitung einer visuellen Repräsentation in der dynamischen Lernumgebung zu einer vertieften Verarbeitung führt (Lenz et al., 2019). Eine schriftliche Begründung der gefundenen Lösung wird gefordert, da eine Verbalisierung der eigenen Lösung eine wichtige Form der aktiven Wissensorganisation ist (Prediger et al., 2021).

Fragestellung

Die Studie untersucht, ob in der dynamischen Lernumgebung mehr strategische Ansätze –als Operationalisierung für ausgelöste kognitive Aktivitäten – generiert werden und, ob dies zu mehr konzeptuellem Wissen führt.

Methode

Zur Überprüfung dieser Fragestellung wurden die verbalen Begründungen beider Gruppen im Posttest entsprechend einem hierarchischen Stufenmodell der Strategien (angelehnt an Boomgaarden et al., 2022) kodiert, welches das Zielwissen in mehrere richtige und falsche Strategien unterteilt. Durch die theoretische Fundierung des Stufenmodells erlaubt die Zuordnung der Schülerlösungen zu einer dieser Strategien den Schluss auf die zugrunde liegenden kognitiven Prozesse. Daraufhin wurden die Lernprozesse der dynamischen Gruppe im Hinblick auf die kognitiven Aktivitäten fokussiert. Die erstellten Lösungen innerhalb der Lernumgebung wurden

ebenfalls entsprechend dem Stufenmodell kodiert und im Anschluss daran für jeden Lernenden die maximal erreichte Stufe im Lernprozess bestimmt.

Ergebnisse und ihre Bedeutung

In einer ersten Pilotierung mit 92 Lernenden zeigte sich ein deskriptiver Trend, dass Lernende in der dynamischen Variante im Posttest besser abschnitten als die statische Gruppe; der Gruppenunterschied war jedoch nicht signifikant ($\chi^2(2, N = 92) = 1.65, p = .44$). Eine Analyse der Prozessdaten mittels „Sequence Clustering“ (zur Anwendung der Methode in den empirischen Bildungswissenschaften, siehe Heitzmann et al., 2022) soll hier weiterführende Einblicke in Unterschiede in den Lernprozessen zwischen den Gruppen geben. Zudem gab es einen moderaten Zusammenhang von $\rho = .48$, $p < .001$ zwischen maximal erreichter Stufe in der dynamischen Lernphase und der erreichten Stufe im Posttest. Dieses Ergebnis zeigt, dass die strategischen Ansätze in der dynamischen Lernumgebung in Zusammenhang mit dem Abschneiden im Posttest stehen. Dieser Zusammenhang dient für die geplante Mediationsanalyse, um zu untersuchen, ob Variationen in der dynamischen Lernphase mehr strategische Ansätze auslösen und, ob diese zu mehr konzeptuellem Wissen führen.

Testing-Effekt und Transfer im mathematischen Lernen: besser gemerkt - aber auch besser angewendet?

Moritz Wußing[1], Roland Grabner[2], Henrik Saalbach[1]
[1]Universität Leipzig, Deutschland; [2]Universität Graz, Österreich

Für das langfristige Behalten von Informationen scheint es förderlicher zu sein, wenn nach einer initialen Lernphase eine Testung statt einer zeitlich äquivalenten erneuten Präsentation des Lernmaterials ohne Übungstest erfolgt (z.B. Carrier & Pashler, 1992; Pyc & Rawson, 2010). Dieses Phänomen wird als „Testing Effekt“ bezeichnet. Für die pädagogische Wirksamkeit dieses Effekts ist nicht nur das verbesserte Behalten der Informationen relevant, wünschenswert wäre auch eine bessere Anwendbarkeit bzw. Transferfähigkeit. In einer Metaanalyse zeigten (Pan & Rickard, 2018), dass ein solcher Transfer ebenfalls durch eine Testung gefördert wird (d = 0.40). Ein genauerer Blick offenbart jedoch, dass hier sehr verschiedene Arten des Transfers betrachtet wurden. Studien, die lediglich einen Nahtransfer untersuchten, z.B. auf ein anderes Testformat (Karpicke & Blunt, 2011) oder eine Umordnung von Reihenfolgen der Stimuli (Carpenter et al., 2006), konnten einen Transfer des TE nachweisen. Von den Studien, die sich auf einen Ferntransfer bezogen (im Sinne von Anwendung oder Problemlösung auf Basis der Informationen), zeigten jedoch nur einige einen signifikanten Vorteil des Testens (Rohrer et al., 2010), andere hingegen nicht (van Gog & Kester, 2012; Wooldridge et al., 2014).

Somit gibt es für die entscheidende Frage nach der durch das Testen geförderten Transferierbarkeit noch keine eindeutige Antwort.

Ziel unserer Arbeit war es daher, den Transfer des Testing-Effekts zu untersuchen, speziell im Bereich des mathematischen Lernens. Im ersten der Experimente untersuchten wir dafür den Nahtransfer auf ein anderes Testformat, im zweiten Experiment untersuchten wir Ferntransfer in Form von Anwendungsaufgaben.

In Experiment 1 lernten 78 Studierende 12 artifizielle Gleichungen der Form 16 □ 8 = 48, die mit dem fiktiven Operator □ ohne eine innere Logik zusammengefügt wurden. Nach einer initialen Lernphase wurde in einem within-subject-design die Hälfte der Gleichungen mit einem cued-recall Übungstest mit Feedback, die andere Hälfte mit einer zeitlich äquivalenten Wiederholungsphase gefestigt. Die Zuteilung der Gleichungen auf die beiden Bedingungen erfolgte randomisiert.

In Experiment 2 lernten 120 Studierende mathematische Fachbegriffe und Zusammenhänge zu den Themen Aussagenlogik und Abbildungen aus dafür konzipierten kurzen Fachtexten. Auch hier wurde, nach einer initialen Lernphase, in einem within-subject-design einer der Texte mit einem cued-recall Übungstest mit Feedback, der andere mit einer Wiederholungsphase trainiert. Auch hier erfolgte die Zuteilung der beiden Texte zu den Bedingungen randomisiert.

In beiden Experimenten konnte in einem Posttest (ohne Feedback) nach einer Woche ein signifikanter TE für einen Test, der lediglich eine Reproduktion der Inhalte forderte (cued-recall-format) festgestellt werden (Experiment 1: $p = .022$, $d = 0.26$; Experiment 2: $p < .001$, $d = 0.79$).

Direkt im Anschluss an diesen ersten Posttest, folgte in beiden Experimenten ein Transfertest. In Experiment 1 wurde im Sinne des Nahtransfers nun ein free-recall statt des cued-recalls aus dem Übungstest gefordert. Hier zeigte sich kein signifikanter Vorteil der Items aus der Testing-Bedingung mehr ($p = .352$, $d = 0.11$). In Experiment zwei bearbeiteten die Teilnehmer einen Test zum Ferntransfer mit Anwendungsaufgaben, in dem die gelernten Begriffe inhaltlich genutzt werden mussten. Auch hier zeigte sich kein signifikanter Vorteil der Begriffe aus der Testing-Bedingung mehr ($p = .179$, $\eta_p^2 = .013$).

Die bessere Transferierbarkeit der Kenntnisse aus den Testing-Bedingungen konnte demnach sowohl in einer Nahtransfer- als auch Ferntransfer-Bedingung nicht gezeigt werden. Unser Ergebnis zum Nahtransfer steht im Widerspruch zu bisherigen Ergebnissen (Karpicke & Blunt, 2011). Für den Ferntransfer sind unsere Ergebnisse ein weiterer Hinweis darauf, dass ein solcher Transfer nicht als selbstverständlich angenommen werden kann (Wooldridge et al., 2014). Neben der pädagogischen Relevanz dieser Ergebnisse gilt es, diese im Licht der kognitiven Theorien des TEs (zum Beispiel transfer-appropriate-processing, (Morris et al., 1977)oder einer zu hohen Elementinteraktivität bei Anwendungsaufgaben (Leahy et al., 2015)) zu analysieren und eine mögliche Sonderstellung des mathematischen Materials zu erwägen.

G-04: Instruktionale Methoden

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S06 S02 B06
Chair der Sitzung: Laura Gehreke

Wirksamkeit von Peer-Mentoring in der Studieneingangsphase: ein systematisches Review

Laura Gehreke, Simone Kauffeld, Hannes Schilling
Technische Universität Braunschweig, Germany

Studienabbrüche sind keine Seltenheit. Zahlen belegen, dass 25% aller Bachelorstudierenden ihr Studium abbrechen (Heublein & Schmelzer, 2018). In Hinblick auf den Studienverlauf lässt sich ein hoher Anteil der Studienabbrüche in der Studieneingangsphase (1.-3. Semester) feststellen (Heublein & Schmelzer, 2018), was längerfristige psychische Folgen auf individueller Ebene, z.B. eine geringere Berufs- und Lebenszufriedenheit (Klein et al., 2021) sowie erhebliche finanzielle Folgen für die Hochschulen haben kann (Madner, 2015).

Um dem Problem zu begegnen, gibt es zahlreiche Maßnahmen zur Unterstützung von Studierenden in der Studieneingangsphase. Eine der häufig eingesetzten Maßnahmen ist das Peer-Mentoring. Es wird definiert als „eine Interventionsstrategie, bei der ein oder mehrere Studierende (d.h. Mentees) mit einem erfahreneren Studierenden (d.h. Peer-Mentor) zusammengebracht werden, der den Mentee(s) sowohl praktische Anleitung als auch soziale Unterstützung bietet" (Yomtov et al., 2017, S.26). " Es soll dabei helfen, die akademische und soziale Integration von Studienanfänger*innen zu fördern (Flores & Estudillo, 2018). Nach Tinto (1993) sind die akademische und soziale Integration entscheidende Faktoren, die die Entscheidung, an der Universität zu bleiben, bestimmen. In diesem Zusammenhang wird die akademische Integration formell durch die akademischen Leistungen der Studierenden und informell durch die Interaktion mit den Lehrkräften und die Wahrnehmung positiver Erfahrungen definiert. Bei der sozialen Integration spielen außeruniversitäre Aktivitäten und die Unterstützung durch Gleichaltrige eine entscheidende Rolle. Im Allgemeinen ist es wahrscheinlicher, dass Studierende, die sich stärker in die Universität integriert fühlen, ihr Studium abschließen. Somit kommt Peer-Mentoring insbesondere bei der sozialen Integration eine entscheidende Funktion zu.

Umfassende Angebote zur Unterstützung von Studierenden werden von Hochschulen eingesetzt, jedoch gibt es kaum fundierte Aussagen zur Wirksamkeit dieser Maßnahmen (Schubarth et al., 2018). Ziel dieses Reviews ist eine systematische Analyse bisheriger Studien im Hinblick auf die Wirksamkeit von Peer-Mentoring als Unterstützungsmaßnahme in der Studieneingangsphase. Dabei soll die Wirksamkeit an den zentralen Variablen Abbruchintention, Zugehörigkeitsgefühl zur Universität und akademische Leistung gemessen werden. Da Frauen beim Übergang von der Schule zur Universität mehr Stress empfinden und dadurch stärker psychisch beeinträchtigt werden (Conley et al., 2014), soll zudem betrachtet werden, ob sich die Effektivität von Peer-Mentoring zwischen den Geschlechtern unterscheidet. Zudem mussten, insbesondere aufgrund der COVID-19 Pandemie, sämtliche Angebote digital angepasst werden (Breitenbach, 2021), sodass eine weitere Fragestellung mögliche Unterschiede in der Effektivität von digitalen vs. nicht digitalen Peer-Mentoring betrachtet.

Bei der Erstellung des systematischen Reviews wird nach den PRISMA-Guidelines vorgegangen (Page et al., 2021). Zu Beginn des Suchprozesses wurde eine Recherche bei Google-Scholar durchgeführt, um bereits bestehende Reviews und zentrale Suchbegriffe zu identifizieren und anschließend bei SCOPUS einen Suchstring zu entwickeln. Der Suchstring wurde daraufhin in drei

verschiedenen Datenbanken zur Recherche eingesetzt: SCOPUS, Web of Science und Ebscohost. Die Suchergebnisse beinhalten nur Artikel nach 1990, da sich in den 1990er Jahren Peer-Mentoring nach heutigen Verständnis etablierte (Unternehmensberatung, 2009). Die Recherche ergab insgesamt 25 562 Artikel. Nach Entfernung der Duplikate blieben 17 801 Artikel für die weitere Auswertung, deren Abstracts von zwei unabhängigen Personen hinsichtlich der Einschlusskriterien geratet wurden. Insgesamt werden ca. 40 Studien in diesem systematischen Review inkludiert. Ergebnisse zeigen, dass Peer-Mentoring eine wirksame Unterstützungmaßnahme für Studienanfänger*innen ist. Insbesondere die soziale Komponente des Peer-Mentorings wies vorwiegend bei Frauen eine hohe Wirksamkeit auf. Es konnte festgestellt werden, dass digitales Peer-Mentoring durchaus eine wirksame Alternative zum herkömmlichen Peer-Mentoring ist. Peer-Mentoring ist somit ein wirksames Unterstützungsangebot, welches aufgrund seiner Vielfältigkeit und unterschiedlichen Komponenten (soziale, z.B. in Form von gemeinsamen außeruniversitären Aktivitäten und akademisch, z.B. Tipps für eine Literaturrecherche) auf unterschiedliche Bedürfnisse angepasst und eingesetzt werden kann. Die Ergebnisse werden vor dem Hintergrund eines vermehrten Einsatzes digitaler Formate diskutiert.

Problem Posing und Motivation - Steigert das Entwickeln eigener Mathematikaufgaben die Selbstwirksamkeit und die Wertüberzeugungen vor und nach der Aufgabenbearbeitung?

Janina Krawitz, Luisa Hartmann, Stanislaw Schukajlow
Universität Münster, Germany

Motivationale Variablen wie Erwartungen und Wertüberzeugungen tragen entscheidend zum erfolgreichen Lernen von Mathematik bei (Schukajlow et al., 2017). Nach Erwartungs-Wert Modellen beeinflussen Erfolgserwartungen und Wertüberzeugungen Leistungen und Bildungsentscheidungen von Lernenden (Eccles & Wigfield, 2020). Demzufolge sollten Interventionen zur Förderung mathematischer Kompetenzen darauf abzielen, Erfolgserwartungen und Wertüberzeugungen positiv zu beeinflussen. Darauf basierend ist das Ziel der vorliegenden Studie, die Wirkung einer Intervention zum Problem Posing auf die vor und nach der Aufgabenbearbeitung berichtete Selbstwirksamkeit und die Wertüberzeugungen zu untersuchen und damit einen Beitrag zur Analyse motivationssteigernder Faktoren zu leisten.

Dem Problem Posing – dem Entwickeln eigener Probleme – wird ein großes Potenzial für die Förderung motivationaler Variablen des Lernens von Mathematik zugesprochen (Cai & Leikin, 2020). Insbesondere die Selbstwirksamkeit und die Wertüberzeugungen könnten vom Problem Posing profitieren. Diese Konstrukte werden maßgeblich durch persönliche Zielsetzungen beeinflusst (Eccles & Wigfield, 2020). Problem Posing bietet die Möglichkeit, durch das Entwickeln von Aufgaben eigene Ziele zu setzen. Erste Ergebnisse aus Studien mit längeren Interventionen weisen auf einen positiven Effekt des Problem Posings auf die Selbstwirksamkeit hin (Akay & Boz, 2010; Chen et al., 2015). Allerdings ist bislang ungeklärt, inwiefern kurzfristige Problem Posing Interventionen (ähnlich wie andere Interventionsansätze (Brisson et al., 2017)) eine motivationsförderliche Wirkung in Bezug auf die Selbstwirksamkeit und die Wertüberzeugungen bezüglich der Aufgabenbearbeitung haben und ob diese – sofern vorhanden – auch nach dem Lösen der Aufgaben bestehen bleibt.

Hypothesen:

H1: Das Entwickeln eigener Mathematikaufgaben wirkt sich positiv auf die Selbstwirksamkeit und die Wertüberzeugungen vor der Aufgabenbearbeitung aus.

H2: Das Entwickeln eigener Mathematikaufgaben wirkt sich positiv auf die Selbstwirksamkeit und die Wertüberzeugungen nach der Aufgabenbearbeitung aus. Der positive Effekt wird über die Einschätzungen vor der Aufgabenbearbeitung vermittelt.

Die Stichprobe umfasst 210 Neunt- und Zehntklässlern (N_{EG} = 127, N_{KG1} = 41, N_{KG2} = 42; 43,8% weiblich; *M* = 15,29 Jahre). Lernende innerhalb einer Klasse wurden randomisiert den Untersuchungsbedingungen zugewiesen. Die Experimentalgruppe (EG) erhielt ein Testheft mit sechs realitätsbezogenen Situationsbeschreibungen, zu denen mathematische Aufgaben entwickelt und anschließend bearbeitet werden sollten. Die Kontrollgruppen (KG1, KG2) erhielten analoge Testhefte mit vorgegebenen Aufgaben. Als vorgegebene Aufgaben wurden die häufigsten Aufgaben ausgewählt, die in einer Vorstudie zu denselben realitätsbezogenen Situationsbeschreibungen entwickelt wurden (Hartmann et al., 2021). Die Selbstwirksamkeit und die Wertüberzeugungen bezüglich der Aufgabenbearbeitung, wurde situationsspezifisch mittels Fragebögen vor und nach der Bearbeitung jeder Aufgabe erfasst. Die Lernenden beantworteten 24 Items (12 vor und 12 nach den Aufgabenbearbeitungen) bestehend aus einer realitätsbezogenen Situationsbeschreibung, einer Aufgabe und einer Aussage zur Selbstwirksamkeit bzw. Wertüberzeugungen (5-stufige Likert-Skala; interne Konsistenz >.83).

Wie erwartet wirkte sich Problem Posing positiv auf die Selbstwirksamkeit der Lernenden aus (EG vs. KG1: β = .540 *p* <. 001; EG vs. KG2: β = .732 *p* < .001). Allerdings hatte Problem Posing entgegen der Erwartungen keinen positiven Effekt auf die Wertüberzeugungen (EG vs. KG1: β = .265 *p* = .071; EG vs. KG2: β = .085 *p* = .317). Der positive Effekt auf die Selbstwirksamkeit blieb auch nach der Aufgabenbearbeitung bestehen (EG vs. KG1: β = .450 *p* < .001; EG vs. KG2: β = .479 *p* < .001) und wurde über die Einschätzung der Selbstwirksamkeit vor der der Aufgabenbearbeitung vermittelt (EG vs. KG1: β = .444 *p* < .001; EG vs. KG2: β = .602 *p* < .001).

Die Ergebnisse ergänzen den Forschungsstand, indem sie die positive Wirkung von kurzfristigen Interventionen zum Problem Posing auf die Selbstwirksamkeit bestätigen und darauf hinweisen, dass individuelle Zielsetzung motivationssteigernde Wirkungen haben können (Eccles & Wigfield, 2020). Als praktische Implikation folgt, dass Problem Posing Aktivitäten im Unterricht zur Entwicklung der Selbstwirksamkeit beitragen können (Chen et al., 2015). Allerdings scheinen umfangreichere Problem Posing Interventionen nötig zu sein, um auch die Wertüberzeugungen zu steigern und beide Motivationskomponenten zu adressieren.

Produktive Nutzung von Fehlern: Ein systematisches Literaturreview

Sonja Dieterich, Stefan Rumann, Marc Rodemer
Universität Duisburg-Essen, Germany

Theoretischer Hintergrund

Aus theoretischer Perspektive ist es keine neue Erkenntnis, dass Fehler lernförderlich sein können und dabei Ausgangspunkt von Lernprozessen verkörpern. Bereits vor 23 Jahren konstatieren Oser et al. (1999) mit ihrer *Theorie des negativen Wissens* eine Lernförderlichkeit ausgehend von Fehlern,

um *negatives Wissen* zu generieren. Dabei können Fehler als Demonstrationsmittel dienen, wie etwas nicht ist oder nicht funktioniert (Oser et al., 1999). Des Weiteren schreiben die drei Dimensionen der Fehleroffenheit nach Weingardt (2004) der Integration von Fehlern im Lernprozess ein positives Potenzial zu. Beim Lernen aus Fehlern dominiert die Verwendung von *erroneous examples*, die in das Feld des *example-based learnings* einzuordnen sind (Renkl, 2014). Worked examples sind durch eine Aufgabenstellung mit anschließender Schritt für Schritt Darstellung der Lösung aufgebaut (Wittwer & Renkl, 2010). Speziell *erroneous examples* präsentieren fehlerhafte Aufgaben, Beispiele oder Lösungen, in denen Fehler beispielsweise identifiziert und korrigiert werden sollen (Große & Renkl, 2007). Das Lernpotenzial ausgehend von *erroneous examples* wird im Bildungskontext derzeit von der vorwiegend fehlervermeidenden Fehlerkultur ausgebremst.

Fragestellung/ Zielsetzung

Das Ziel dieser Arbeit war ein systematisches Literaturreview zur Lernförderlichkeit von Fehlern in Lehr-/Lernkontexten. Dieses Review folgt dem Verständnis von Fehlern als *erroneous examples*. Ein besonderes Augenmerk lag auf der produktiven Nutzung von Fehlern im naturwissenschaftlichen Unterricht, da gerade im Chemie-, Biologie- und Physikunterricht dieses Lernpotenzial bisher wenig genutzt wird. Es ist der Frage nachgegangen worden, wie Fehler im Lernmaterial implementiert werden müssen, um lernförderlich zu sein.

Methode

Zur Orientierung für die systematische Literaturrecherche dient der methodische Leitfaden für systematische Übersichten von Alexander (2020). Das methodische Vorgehen ist geprägt von zwei explorativen Suchphasen zu Beginn, woraus fünfzehn Suchbegriffe generiert und festgelegt wurden. Mit diesen Suchbegriffen ist zum einen eine elektronische Suche in den Datenbanken *Web of Science, ERIC, PubPsych* und *Pedocs* und zum anderen eine manuelle Suche in Referenzlisten und anderen Quellen durchgeführt worden. Daraus ergab sich eine Gesamtzahl von 604 Publikationen. Nach Entfernung der Duplikate sind die resultierenden 437 Publikationen mithilfe von Ein- und Ausschlusskriterien in einem ersten inhaltlichen Screening, wodurch die Publikationen auf 88 reduziert wurden, und einem zweiten inhaltlich und methodischem Screening bezüglich des Einschlusses in dieses Review analysiert worden. Es wurden nur Studien im Prä-/Post-Test Design mit Kontrollgruppe oder Vergleichsgruppe aufgenommen. Aus der systematischen Suche mit anschließender Kategorisierung und Systematisierung resultieren aus den 30 für dieses Review eingeschlossenen Publikationen insgesamt 36 Studien.

Ergebnisse und ihre Bedeutung

Dieses systematische Literaturreview zur produktiven Nutzung von Fehlern kann darlegen, dass einige empirische Publikationen zum Lernen aus *erroneous examples* vorliegen. Die Ergebnisse zeigen für 25 der 36 inkludierten Studien eine Lernförderlichkeit von *erroneous examples*, welche sich positiv auf das Lernen auswirken und folglich Lernpotenzial aufweisen. Diese Lernförderlichkeit tritt sowohl bei Schülerinnen und Schülern (56 % der Studien, n= 14) als auch bei Studierenden und Auszubildenen (44 % der Studien, n= 11) auf, sodass ein altersübergreifender lernförderlicher Effekt von *erroneous examples* in den Studien sichtbar ist.

Insgesamt zeigen die Ergebnisse lernförderliche Effekte, wenn Lernende in *erroneous examples* Fehler z.B. identifizieren, erklären, korrigieren und/oder reflektieren müssen. Dabei unterstützen elaboriertes, teils korrigierendes, Feedback und Selbsterklärungsprompts effektiv den Lernprozess. Zudem beeinflussen weitere Faktoren wie Vorwissen und Cognitive load die Lernförderlichkeit beim Lernen mit *erroneous examples*, wobei starke Abhängigkeiten zwischen dem zu erlernenden, domänenspezifischen Lerngegenstand und der Fehlerinstruktion im Lernmaterial bestehen.

Während eine lernförderliche Implementation von *erroneous examples* in der Mathematik und der Medizin dominierend vorzufinden ist, zeigt dieses Review für den naturwissenschaftlichen Unterricht mit n= 4 Studien kaum eine Verwendung von *erroneous examples* trotz der lernförderlichen Wirkung. Folglich gilt es, in zukünftigen Forschungsvorhaben ein lernförderliches Lernmaterial mit entsprechender Fehlerinstruktion für den naturwissenschaftlichen Unterricht zu entwickeln und empirischen Studien zu unterziehen.

Feedback-Memory: Ein Konzept zur Erhöhung des wahrgenommenen Lehrkraftlobs und der sozialen Akzeptanz?

Markus Spilles, Christian Huber, Nicolay Philipp, Hank Corinna
Bergische Universität Wuppertal

Theoretischer Hintergrund

Internationale Studien lassen vermuten, dass Lehrkräfte durch ihr Feedbackverhalten einen Einfluss auf die soziale Akzeptanz von Schülerinnen und Schülern in inklusiven Schulklassen nehmen könnten (bspw. Huber et al., 2018; Hendrickx et al., 2017; Spilles et al., angenommen). Hintergrund hiervon ist die Annahme, dass Schulkinder insbesondere das Verhalten ihrer Klassenlehrkraft gegenüber einem Kind als Referenz für dessen soziale Akzeptanz nutzen.

Interventionsansätze zur Verbesserung der sozialen Akzeptanz durch eine Modifikation des Lehrkraftfeedbacks wurden bislang noch nicht evaluiert. Im aktuellen Beitrag wird daher ein Konzept zur Erhöhung des durch die Mitschülerinnen und Mitschüler wahrgenommenen Lehrkraftlobs vorgestellt und evaluiert (Feedback-Memory).

Fragestellung

Im Beitrag werden zwei Fragestellungen beleuchtet:

1. Erhöht sich durch die Durchführung des Feedback-Memory-Konzepts das durch die Mitschülerinnen und Mitschüler wahrgenommene Lehrkraftlob im Hinblick auf Kinder, die von ihrer Lehrkraft eher selten im Unterricht gelobt werden?

2. Verbessert sich über diesen Weg auch die soziale Akzeptanz dieser Zielkinder durch ihre Mitschülerinnen und Mitschüler?

Methode

Die Untersuchung wurde im Rahmen eines quasi-experimentellen Wartekontrollgruppendesigns umgesetzt, bei dem in ca. 10 dritten und vierten Klassen das Feedback-Memory-Konzept durchgeführt wurde, ca. 10 weitere Klassen fungierten als Kontrollgruppe. In den Blick genommen

werden nur Kinder, die von ihrer Lehrkraft im Unterricht eher selten gelobt werden. Da die Studie aktuell noch läuft, ist eine finale Beschreibung der Stichprobe leider noch nicht möglich.

Mit Hilfe von Peer-Ratings wurden zu zwei Messzeitpunkten (Prä/Post) das durch die Mitschülerinnen und Mitschüler wahrgenommene Lehrkraftlob sowie die soziale Akzeptanz erfasst.

Die statistische Auswertung erfolgt anhand von zwei Mehrebenenanalysen, bei denen sowohl die Nestung der beiden Messzeitpunkte innerhalb der Individuen als auch die Nestung der Individuen in ihren Klassen berücksichtigt werden.

Ergebnisse

Erste statistische Analysen deuten im Vergleich der unterschiedlichen Entwicklungen von Experimental- und Kontrollgruppe auf eine signifikante Erhöhung des durch die Mitschülerinnen und Mitschüler wahrgenommenen Lehrkraftlobs zugunsten der Feedback-Memory-Gruppe hin. Im Hinblick auf die Veränderung der sozialen Akzeptanz zeigen sich in den ersten Analysen noch keine signifikanten Effekte, auch wenn die deskriptive Statistik auf eine ähnlich positive Entwicklung hindeutet. Einschränkend muss jedoch erwähnt werden, dass zum Zeitpunkt der Beitragseinreichung noch ca. 50 % der Daten fehlen.

Diskussion

Auf Basis der ersten Ergebnisse erscheint das Feedback-Memory-Konzept als ein vielversprechender Ansatz zur Erhöhung des durch die Mitschülerinnen und Mitschüler wahrgenommenen Lehrkraftlobs. Ob auch eine Verbesserung der sozialen Akzeptanz hierdurch möglich ist, müssten künftige Studien mit mehr Teilnehmenden erörtern. Außerdem wäre eine differenzierte Analyse von Wirkfaktoren der Maßnahme wünschenswert.

G-05: Prokrastination und Resilienz

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S06 S03 B06
Chair der Sitzung: Lisa Bäulke

Verändern sich Prokrastinationstendenzen im Verlauf des jungen Erwachsenenalters? Ergebnisse einer Längsschnittstudie zu Entwicklung, Prädiktoren und Auswirkungen von Prokrastination

Lisa Bäulke[1], Brent Roberts[2], Benjamin Nagengast[1], Ulrich Trautwein[1]
[1]Universität Tübingen, Deutschland; [2]University of Illinois Urbana-Champaign

Prokrastination – ein irrationaler Handlungsaufschub, der mit erheblichen negativen Konsequenzen einhergeht – ist ein weit verbreitetes Phänomen, insbesondere im akademischen Kontext (vgl. Steel, 2007). Bis zu 65% der Studierenden möchten ihre Prokrastination verringern (Hayat et al., 2020). Doch können Personen, die Prokrastinationstendenzen aufweisen, diese im Verlauf ihres Lebens überwinden? Obwohl Prokrastination ein dysfunktionales Konstrukt mit bildungsrelevanter Tragweite darstellt, fehlt bisher ein Gesamtbild über die Lebensspanne hinweg.

Prokrastination wurde vorwiegend aus einer sozio-kognitiven Perspektive betrachtet (Sirois & Pychyl, 2013; Steel & König, 2006) und auf Gewissenhaftigkeit und Neurotizismus zurückgeführt (van Eerde, 2003). Jedoch fehlt Wissen über die Kontinuität dieses Konstrukts. Personen könnten im Verlauf ihres Lebens einen besseren Umgang mit herausfordernden Aufgaben erlernen oder durch den Eintritt in den Beruf ihre Prokrastination überwinden. Querschnittlich angelegte Forschung zur Persönlichkeitsentwicklung zeigt beispielsweise, dass Persönlichkeitsmerkmale entwicklungsbedingte Konstrukte sind, mit bemerkenswerten Unterschieden im jungen Erwachsenenalter (18-40 Jahre; Roberts & Mroczek, 2008; Slobodskaya, 2021). Da allerdings großangelegte Längsschnittstudien fehlen, ist bislang unklar, ob und wie sich Prokrastinationstendenzen über das junge Erwachsenenalter hinweg entwickeln.

Um diese Forschungslücke zu adressieren, ist das Ziel der vorliegenden Studie eine umfassende Untersuchung der generellen Entwicklung von Prokrastination. Dabei wird den Fragestellungen nachgegangen, (1) ob und wie sich Prokrastination entwickelt, (2) ob eine Prokrastinationsentwicklung mit Gewissenhaftigkeit und Neurotizismus assoziiert ist, (3) ob eine Prokrastinationsentwicklung mit einem Übergang von der Universität in den Beruf assoziiert ist und (4) welche Konsequenzen Prokrastination für zukünftige Lebensereignisse hat.

Zur Beantwortung der Fragestellungen wurden Daten der ersten acht Erhebungswellen einer fortlaufenden Längsschnittstudie analysiert (TOSCA – Transformation des Sekundarschulsystems und akademische Karrieren, Trautwein et al., 2010). Die erste Erhebung erfolgte im letzten Schuljahr vor dem Abitur, gefolgt von fünf weiteren Erhebungswellen im Zweijahresrhythmus und zwei weiteren Erhebungswellen im Vierjahresrhythmus. Die Studie umfasst folglich eine Zeitspanne von 18 Jahren, der achte Messzeitpunkt erfolgte während der COVID-19 Pandemie. Die Ausgangsstichprobe bestand aus 3023 Schüler:innen (58% weiblich; 42% männlich; Durchschnittsalter = 19.55 Jahre; *SD*= .78) mit einem Dropout von 41% von Erhebungswelle 1 bis 8. Um dies in den Analysen zu berücksichtigen, wurde für die Teilnahmehäufigkeit kontrolliert. Die Daten wurden mittels Mehrebenenanalysen (Messzeitpunkte genestet in Personen) analysiert.

Die Ergebnisse zeigen, dass erhebliche Varianzanteile auf Unterschiede zwischen Personen zurückgehen (ICC = .63, $p < .001$). Ein Unconditional Growth-Modell zeigte eine lineare Abnahme von

Prokrastination über die Zeit hinweg (b = –.15, $p < .001$) mit signifikanter Slope-Varianz zwischen Personen (*Var* = .05, $p < .001$). Um Gewissenhaftigkeit und Neurotizismus als Prädiktoren von Prokrastination untersuchen zu können, wurde zunächst die Faktorstruktur überprüft. Fit Indizes verwiesen auf eine faktorielle Trennbarkeit von Gewissenhaftigkeit, Neurotizismus und Prokrastination. Anschließend wurden Gewissenhaftigkeit und Neurotizismus als Level 2-Prädiktoren in das Wachstumskurvenmodel aufgenommen. Die Ergebnisse zeigten einen negativen Zusammenhang zwischen Gewissenhaftigkeit und Prokrastinationsniveau (b = –.15, $p < .001$) sowie einen positiven Zusammenhang mit Prokrastinationsentwicklung (b = .05, $p < .001$). Neurotizismus war positiv mit Prokrastinationsniveau (b = .11, $p < .001$) und negativ mit Prokrastinationsentwicklung (b = –.02, $p < .01$) assoziiert. Zudem hing Prokrastinationsentwicklung negativ mit einem Transit in den Beruf zusammen ($r = –.11$, $p < .01$). Abschließend zeigten sich insbesondere negative Auswirkungen von Prokrastination auf Beziehungsvariablen und Gesundheit sowie mentale Gesundheit während der Pandemie.

Das Zusammenhangsmuster bestätigt die postulierten Hypothesen, indem es auf eine allgemeine Abnahme von Prokrastination über das junge Erwachsenenalter hinweg deutet. Weiterhin implizieren die Ergebnisse, dass Gewissenhaftigkeit und Neurotizismus Unterschiede in Prokrastinationstendenzen erklären können. Die abnehmende Prokrastinationsentwicklung scheint mit einem Übergang in den Beruf zusammenzuhängen. Die Ergebnisse weisen außerdem daraufhin, dass sich Prokrastination langfristig nachteilig auf Beziehungsvariablen und Gesundheit auswirkt. Zusammenfassend untersuchte die vorliegende Studie erstmals die Kontinuität von Prokrastinationstendenzen aus einer Entwicklungsperspektive. Dadurch konnten Erkenntnisse zu Entstehung und Folgen dieses dysfunktionalen Verhaltens gewonnen werden.

Prokrastination und das schlechte Gewissen - Gruppenunterschiede im Erleben und Auftreten von Prokrastination

Julia Bobe, Katrin Klingsieck
Universität Paderborn, Deutschland

Akademische Prokrastination bezeichnet ein unter Studierenden weit verbreitetes Phänomen, bei dem eine intendierte Aufgabe, aufgrund von kurzfristiger Stimmungsregulation (Sirois & Pychyl, 2013), vermieden und auf einen späteren Zeitpunkt aufgeschoben wird (Steel & Klingsieck, 2016). Die Stimmungsverbesserung wird langfristig von negativen Konsequenzen überwogen (Klingsieck, 2013) und geht mit einem subjektiven Unwohlsein, ähnlich eines schlechten Gewissens, einher (Solomon & Rothblum, 1984).

Wenig untersucht ist, inwieweit sich Studierende bezüglich dieses subjektiven Unwohlseins voneinander unterscheiden und in welchem Zusammenhang dies mit der Häufigkeit von Prokrastination steht. Erkenntnisse dazu würden zum theoretischen Verständnis des Teufelskreises von Prokrastination (vgl. Wäschle et al., 2014) und bei der Entwicklung von Interventionen beitragen.

Eine aktuelle Studie zeigt einen geringen Zusammenhang zwischen dem subjektiven Unwohlsein und der Häufigkeit von Prokrastination (Bobe et al., 2022). Theoretische Überlegungen legen jedoch gruppenspezifische Unterschiede nahe (Milgram, 1993). Das antizipierte subjektive Unwohlsein könnte mit einer geringeren Häufigkeit von Prokrastination einhergehen. Andererseits könnte es

geringere Selbstwirksamkeit und Aufgabenaversivität begünstigen und Prokrastination steigern (Wäschle et al., 2014). Zuletzt könnte es keinen Zusammenhang geben (Milgram, 1992).

Fragestellung

In diesem Beitrag werden gruppenspezifische Unterschiede in Bezug auf das subjektive Unwohlsein während einer Prokrastinationsepisode und dem Zusammenhang mit der Häufigkeit von Prokrastination untersucht.

Methode

N = 891 Studierende ($n_{weiblich}$ = 76.8 %; *M* = 24.17 Jahre) wurden online befragt. Akademische Prokrastination wurde mittels der Subskalen *Aufschub* (Ω = .89) und *subjektives Unwohlsein* (Ω = .86) der Behavioural-and-Emotional-Academic-Procrastination-Scale (BEPS; Bobe et al., 2022) erfasst. Zudem wurden Selbstwirksamkeit (Respondek et al., 2017; Ω = .84), Studienzufriedenheit (Schiefele & Jacob-Ebbinghaus, 2006; Ω = .87) und akademische Emotionen (Pekrun et al., 2011; negative Emotionen, Ω = .77; positive Emotionen, Ω = .67) erfasst. Die Daten wurden mittels hierarchischer Clusteranalyse und einfaktorieller Varianzanalyse ausgewertet.

Ergebnisse

Es finden sich drei Cluster (s. Tabelle 1). Studierende in Cluster 1 berichten hohes subjektives Unwohlsein und niedrige Häufigkeit von Prokrastination. Studierende in Cluster 2 berichten niedriges subjektives Unwohlsein und niedrige Häufigkeit von Prokrastination. Studierende in Cluster 3 berichten hohes subjektives Unwohlsein und hohe Häufigkeit von Prokrastination. Eine ANOVA ergab signifikante Mittelwertsunterschiede zwischen den Clustern (Aufschub, $F = 523.96$, $p < .01$; subjektives Unwohlsein, $F = 366.76$, $p < .01$).

Tabelle 1

Mittelwerte und Standardabweichungen der erfassten Variablen.

	Cluster 1	Cluster 2	Cluster 3
N	361 (40.51%)	417 (46.80%)	113 (12.68%)
Aufschub	2.24 (.55)	3.13 (.71)	4.36 (.52)
Subjektives Unwohlsein	4.11 (.63)	3.07 (.79)	4.72 (.34)
Selbstwirksamkeit	3.89 (.75)	3.89 (.69)	3.53 (.89)
Studienzufriedenheit	4.02 (.85)	4.04 (.81)	3.51 (.96)
neg. akad. Emotionen	2.47 (.79)	2.41 (.74)	3.05 (.96)
pos. akad. Emotionen	3.69 (.73)	3.69 (.64)	3.24 (.83)

Eine weitere ANOVA zeigt signifikante Mittelwertsunterschiede zwischen den Clustern für Selbstwirksamkeit ($F = 11.59$, $p < .01$), Studienzufriedenheit ($F = 18.63$, $p < .01$), sowie positive ($F = 19.05$, $p < .01$) und negative akademische Emotionen ($F = 30.51$, $p < .01$). Ein Scheffé Post Hoc Test ergab, dass Cluster 3 über alle Variablen hinweg einen signifikant geringeren positiven Affekt berichten als Cluster 1 und 2.

Diskussion

Für Gruppe 1 ist subjektives Unwohlsein offenbar funktional, da diese über die niedrigste Häufigkeit von Prokrastination berichtet. Für Gruppe 2 scheinen Prokrastination und subjektives Unwohlsein keine maßgebliche Rolle zu spielen. Gruppe 3 unterscheidet sich von Gruppe 1 nicht so deutlich in der Intensität subjektiven Unwohlseins, jedoch in der berichteten Häufigkeit von Prokrastination. Studierende, die allgemein über hohe negative Affektivität berichten, prokrastinieren am häufigsten und fühlen sich dabei am schlechtesten. Diese Interpretation findet sich auch in Befunden wieder, die negative Zusammenhänge zwischen Prokrastination und der Bewusstmachung negativer Konsequenzen des eigenen Verhaltens (Sirois, 2014) sowie Teufelskreise aus niedriger Selbstwirksamkeit und Prokrastination (Wäschle et al., 2014) zeigen. Zukünftig soll geprüft werden, inwieweit sich die Gruppen auch in der *zukünftigen* Auftretenswahrscheinlichkeit von Prokrastination voneinander unterscheiden.

Wer prokrastiniert am meisten – die Ängstlichen oder die Hoffnungsvollen? Latente Emotionsprofile Studierender und ihre Zusammenhänge mit akademischer Prokrastination

Christopher Kafui Gadosey[1], Derya Turhan[1], Theresa Schnettler[1], Anne Scheunemann[2], Lisa Bäulke[3], Daniel O. Thies[2], Markus Dresel[3], Stefan Fries[4], Detlev Leutner[5], Carola Grunschel[1]
[1]Westfälische Wilhelms-Universität Münster, Germany; [2]Ruhr-Universität Bochum; [3]Universität Augsburg; [4]Universität Bielefeld; [5]Universität Duisburg-Essen

Empirische variablenzentrierte Studien zeigen Emotionen als wichtige Korrelate der Prokrastination (Behnagh & Ferrari, 2022; Steel, 2007) und Prokrastination wird sogar als Strategie zur Emotionsregulation betrachtet (Pychyl & Sirois, 2016). Bisherige Studien berücksichtigten allerdings nicht, dass verschiedene Emotionen gleichzeitig in einer Person vorkommen und verschiedene Emotionskonstellationen unterschiedlich mit Prokrastination zusammenhängen können (Gadosey et al., 2021). Einige Studien identifizierten Profile mit mehreren Emotionen und zeigten, dass sich diese Profile auch hinsichtlich Studienmotivation, Engagement und Leistung unterscheiden (z. B. Ganotice et al., 2016; Mata et al., 2020). Zusammenhänge zwischen Profilen von Emotionen und dem Prokrastinationverhalten der Studierenden wurden bisher nicht untersucht. Diese Untersuchung könnte jedoch differenzierte Erkenntnisse darüber liefern, wie unterschiedliche Konstellationen positiver und negativer Emotionen mit Prokrastination assoziert sind. Diese Erkenntnise werden wiederum für die Gestaltung zugeschnittener Präventions- und Interventionsmaßnahmen benötigt.

Im Hinblick auf bevorstehende Prüfungen erleben Studierende insbesondere Angst (negative Emotion) und Hoffnung (positive Emotion; Pekrun et al., 2004). Während Angst positiv mit Prokrastination zusammenhängt (Steel, 2007), steht Hoffnung in einem negativen Zusammenhang mit Prokrastination (Alexander & Onwuegbuzie, 2007; Gadosey et al., 2021). Interessanterweise haben Gadosey et al. (2021) gezeigt, dass prüfungsbegzogene Angst und Hoffnung nicht nur gemeinsam auftreten, sondern auch der negative Zusammenhang zwischen Hoffnung und Prokrastination bei höherer Angst schwächer wird. Demzufolge weisen diese Befunde darauf hin, dass unterschiedliche emotionale Profile von Angst und Hoffnung unterschiedlich mit Prokrastinationsverhalten zusammenhängen können (Gadosey et al., 2021). Unklar ist jedoch, ob sich aus diesen beiden häufig gemeinsam auftretenden Emotionen unterschiedliche Profile ableiten lassen und wie diese unterschiedlichen Profile mit akademischer Prokrastination zusammenhängen. Daher untersuchte die vorliegende Studie mittels eines personenzentrierten Ansatzes, ob Studierende

in unterschiedliche prüfungsbezogene Emotionsprofile gruppiert werden können und ob sich diese Gruppen in Bezug auf ihre Tendenzen zur akademischen Prokrastination unterscheiden.

Zur Überprüfung der Fragestellungen wurden Daten einer Studie (*N* = 637) verwendet, die zum Ende der Vorlesungszeit des Wintersemesters 2020/2021 an drei deutschen Universitäten erhoben wurden. Studierende waren in verschiedenen Studienfächern immatrikuliert. Sie beantworteten Fragen zur prüfungsbezogenen Angst (McDonald's ω = .79) und prüfungsbezogenen Hoffnung (McDonald's ω = .88; Pekrun et al., 2011). Die Tendenzen zur akademischen Prokrastination wurden mit den deutschen Versionen der an das Studium angepassten Tuckman Procrastination Scale (TPS-d; McDonald's ω = .96; Stöber & Joormann, 2001) und der Behavioral and Emotional Academic Procrastination Scale (BEPS, McDonald's ω = .91 für Subskala Aufschub und ω =.90 für Subskala Subjektives Unbehagen; Bobe et al., 2022) erfasst, um differenzierte Einblicke in das Prokrastinationsverhalten der Studierenden zu erhalten. Latente Profilanalysen (LPA) wurden zur Identifikation sowie zur Validierung der Profile mit Mplus 8.7 (Muthén & Muthén, 1998-2017) berechnet. Zur Bestimmung der Anzahl der Profile wurden sowohl statistische als auch inhaltliche Kriterien herangezogen (siehe Ferguson et al., 2020).

Es kristallisierten sich drei Emotionsprofile heraus: „hohe Hoffnung–geringe Angst“ (*n* = 153; 24 %), „moderate Hoffnung und Angst“ (*n* = 337; 53 %) und „hohe Angst–geringe Hoffnung“ (*n* = 147; 23 %). Studierende mit hoher Hoffnung und geringer Angst gaben signifikant niedrigere Werte für die TPS-d sowie die Subskalen Aufschub und subjektives Unbehagen der BEPS an als die anderen beiden Profile. Studierende mit moderater Hoffnung und Angst berichteten signifikant weniger Prokrastination (für TPS-d und die beiden Subskalen der BEPS) als Studierende mit hoher Angst und geringer Hoffnung. Insgesamt berichteten die Studierenden mit hoher Angst und geringer Hoffnung das höchste Maß an Prokrastination.

Zusammenfassend implizieren die Befunde, dass eher das Erleben positiver (Hoffnung) statt negativer Emotionen (Angst) der Tendenz zur akademischen Prokrastination möglicherweise entgegenwirken kann. Daher könnte die Förderung von Hoffnung als positive prüfungsbezogene Emotion für einige Lernende (z.B. Studierende mit hoher Angst) eine sinnvolle Intervention zur Reduktion von Prokrastination sein.

Aussagen über akademische Resilienz auf Grundlage von PISA-Daten: Sind die Befunde verfälscht durch niedrige Testanstrengung?

Ulf Kroehne[1], Nani Teig[2], Frank Goldhammer[1,3]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany; [2]University of Oslo, Norway; [3]Centre for International Student Assessment (ZIB)

Theoretischer Hintergrund: Als Ausdruck akademischer Resilienz (AR) wird eine hohe Kompetenz oder Fähigkeit interpretiert, wenn sie bei Schülerinnen und Schülern aus ungünstigen Verhältnissen beobachtet wird. Ein spezifischer Fokus auf der Gruppe resilienter Schüler ist unter anderem im dem Forschungsinteresse der Identifikation von besonders günstigen individuellen oder institutionellen Bedingungen begründet. Häufig werden dabei Daten aus groß angelegten Bildungsvergleichsstudien genutzt (z.B. Agasisti, 2018), welche hohen methodischen Standards u.a. im Hinblick auf eine Populationsbeschreibung genügen. Teilnehmende Schülerinnen und Schüler bearbeiten dabei oft

sowohl Fragebögen als auch kognitive Testaufgaben unterschiedlicher Domänen in einem Setting, in welchem die individuelle Leistung und Testanstrengung keine direkten Konsequenzen haben (*low stakes assessment*). Niedrige Testanstrengung kann nicht nur über Selbsteinschätzungen (siehe u.a. Ulitzsch et al., 2021), sondern in computer-basierten Assessments auch mit Hilfe von Bearbeitungszeiten als *Rapid Guessing* (RG) beobachtet (z.B. Wise, 2017) und mit Hilfe einer Bearbeitungszeitschwelle operationalisiert werden. Aus dieser Forschung zu RG ist unter anderem bekannt, dass niedrige Testanstrengung nicht nur bei männlichen Schülern, sondern auch bei Schülerinnen und Schülern aus ungünstigem sozio-ökonomischem Hintergrund häufiger auftritt (d.h. vermehrt RG gezeigt wird, vgl. bspw. Rios & Soland, 2022). Damit ergibt sich für die Forschung zur akademischen Resilienz die Fragestellung, in wieweit die Identifikation von leistungsstarken Schülerinnen und Schülern aus ungünstigen Verhältnissen durch niedriger Testanstrengung konfundiert ist.

Fragestellung: In diesem Beitrag wird anhand von Daten der PISA-Studie (2015 und 2018) aus Deutschland und der Schweiz vertiefend untersucht, inwieweit die Identifikation von akademisch resilienten Schülerinnen und Schüler durch niedrige Testanstrengung verfälscht ist. Dazu werden verhaltensbezogene Maße berücksichtigt, die mit Hilfe von Bearbeitungszeitschwellen niedrige Testanstrengung anzeigen (RG als proximales Maß im kognitiven Assessment, und *Rapid Responding, RR,* als distales Maß für das äquivalente Verhalten von *Careless Insufficient Effort Responding* im Fragebogen, vgl. Kroehne et al., 2019).

Methode: Für die Identifikation von AR folgen wir dem definitionsgestützten Ansatz (z.B. Rudd et al., 2021), unter Verwendung des ESCS (*Economic, Social and Cultural Status*) für die Erkennung von nachteiligem sozio-ökonomischen Hintergrund und der *Plausible Values* (für Mathematik, Lesen und Naturwissenschaften) zur Identifikation hoher Kompetenz als Ausdruck erfolgreicher Anpassung (z.B. Ye et al., 2021), vergleichend mit relativen und absoluten Schwellen. Für die Klassifikation von Antworten auf kognitive Aufgaben (RG) werden Bearbeitungszeiten wie von der OECD bereitgestellt verwendet, für die Fragebogenitems (RR) wurden die Bearbeitungszeiten aus den Log-Daten extrahiert und als *Average Answering Time* pro Fragebogenbildschirm und Itembatterie zusammengefasst. Zeitschwellen für 320 kognitive Items (RG) und 20 Itembatterien (RR) wurden mit dem R Paket multimode (Ameijeiras-Alonso et al., 2021) bestimmt, wodurch entsprechend der bimodalen Verteilung schnelles Antwortverhalten mit niedriger Testanstrengung klassifiziert und als *Response Time Effort* (RTE) zusammengefasst werden kann. Proximaler RTE (basierend auf RG) und distaler RTE (basierend auf RR) werden genutzt um AR in Subgruppen mit vergleichbarer Testanstrengung zu ermitteln.

Ergebnisse & Diskussion: Die Ergebnisse bestätigen zunächst, dass proximales RG und distales RR zu 0.30 positiv korreliert sind, sich verhaltensbasierte Maße für niedrige Testanstrengung also sowohl aus Bearbeitungszeiten im kognitiven Assessment als auch aus schnellen Fragebogenantworten bestimmen lassen. Sowohl für relative als auch für absolute Schwellen zur Identifikation von AR zeigt sich darüber hinaus die erwartete Konfundierung mit niedriger Testanstrengung, Schülerinnen und Schüler welche mit hoher Leistung trotz ungünstiger sozio-ökonomischer Bedingungen als AR klassifiziert werden, zeigen weniger häufig RR und RG, bearbeiten die

Aufgaben also mit einer höheren Testanstrengung. Betrachtet man AR jedoch für Gruppen mit vergleichbarem Ausmaß von RR und RG getrennt, lassen sich Prädiktoren akademischer Resilienz unabhängig von der beobachteten Testanstrengung finden. Mit dem entwickelten methodischen Vorgehen lässt sich AR als robust bzgl. niedriger Testanstrengung untersuchen. In der abschließenden Diskussion werden weitergehende Möglichkeiten zur Korrektur für niedrige Testanstrengung diskutiert (vgl. Deribo et al., 2021), die über die Betrachtung bedingter Effekte hinausgehen.

G-06: Unterrichtsgestaltung: Authentisch, individualisiert und bedürfnisorientiert

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S06 S04 B06
Chair der Sitzung: Dennis Hauk

Wie bindet man eher schweigsame Schülerinnen und Schüler in das Unterrichtsgespräch ein? Eine videobasierte Fallstudie

Dennis Hauk, Alexander Gröschner
Friedrich-Schiller-Universität Jena, Germany

Theoretischer Hintergrund

Das Klassengespräch zählt über Länder- und Fachgrenzen hinweg zu den häufigsten Interaktionsformen im Unterricht. Die Erkenntnisse der interaktionsorientierten Unterrichtsforschung zeigen, dass die verbale Beteiligung von Lernenden im Klassengespräch einen positiven Einfluss auf das Lernen und die sozial-emotionale Einbindung der Schüler:innen besitzt (Wegerif, 2007; Lipowsky 2007). Auf Grundlage dieser empirischen Erkenntnisse stellt sich die Frage nach dem aktiven Einbezug eher schweigsamer Schüler:innen in das Unterrichtsgespräch. Die vorliegende Studie untersucht diesbezüglich das Kommunikationsverhalten zwischen Lehrpersonen und schweigsamen Schüler:innen (s. g. *Silent High-Achieving Students, SHA)* im Unterricht. Bei dieser Gruppe handelt es sich um Lernende, die ein hohes kognitives Grundpotential besitzen, jedoch kaum eine aktive verbale Beteiligung im Unterrichtsgespräch anstreben (Sedova & Navratilova, 2020). Bisherige Studien konnten zeigen, dass SHA-Schüler:innen im besonderen Maße durch extrinsische Gesprächsimpulse der Lehrperson (z.B. durch das direkte Ansprechen bzw. Aufrufen) in das Unterrichtsgespräch einbezogen werden (Ing et al, 2015). Es liegen bislang jedoch keine systematischen Befunde darüber vor, wie Schüler:innen auf diese Impulse reagieren bzw. wie sie diese wahrnehmen.

Fragestellung

Vor dem Hintergrund des skizzierten Forschungsdesiderats zielt die vorliegende Studie darauf ab, neue Erkenntnisse über das Interaktionsverhalten und die Möglichkeiten der verbalen Einbindung von SHA-Schüler:innen in das Unterrichtsgespräch zu gewinnen. Folgende Fragestellungen wurden untersucht:

1. Wie laden Lehrpersonen SHA-Schüler:innen dazu ein, sich verbal am Unterrichtsgespräch zu beteiligen?
2. Wie antworten SHA-Schüler:innen auf die Gesprächseinladung der Lehrperson?
3. Wie nehmen SHA-Schüler:innen die Gesprächseinladung der Lehrperson wahr?

Methode

Die Datenerhebung erfolgt als explorative Feldstudie im Mixed-Methods-Design unter Nutzung von videogestützten Feldbeobachtungen, Fragebögen und leitfadengestützten Interviews. An einer freien Ganztagsschule wurden aus zwei Untersuchungsklassen (Klasse A und B) der Sekundarstufe (N = 42 Schüler:innen) n = 6 Schüler:innen (5 weiblich; Altersspanne 12 – 14 Jahre) mittels eines stratifizierten Vorgehens als Fokusgruppe ausgewählt. Für die Kodierung der Videoaufnahmen wurde ein Ratingtool entwickelt, dass die Qualität und Quantität der lehrer- und schülerseitigen Äußerungen erfasst (Howe et al., 2019). Die Kodierung erfolgte durch zwei geschulte Rater und weist eine hohe Beurteilerübereinstimmung auf (α_2 = .80 [.781–.835]). Die Wahrnehmung der Schüler:innen wurde mittels Fragebogen (*Dialogic Teaching Questionnaire* (DTQ-S), Gröschner et al., 2020) mit insgesamt 4 Items (6-Punkt Likert Skala; α = .92) erfasst.

Ergebnisse und Diskussion

Die verhaltensbezogene Analyse der Videodaten deutet einen Zusammenhang zwischen der Qualität der lehrerseitigen Gesprächseinladungen und der Äußerungen der SHA-Schüler:innen an. Die mit qualitätsvolleren Gesprächseinladungen adressierten SHA-Schüler:innen der Klasse A äußerten sich nicht nur länger (Bereich der Wortanzahl in Klasse A von n = 12 – 16 Wörter; in Klasse B von n = 1 – *8* Wörter), sondern auch unter Einbezug weiterführender Begründungen (Anzahl der begründeten Repliken in Klasse A n = 5; im Vergleich zu n = 0 in Klasse B). Zudem nehmen die SHA-Schüler:innen in Klasse A (M = 5.00; SD = 0.61) eine höhere Qualität der Gesprächseinladungen wahr als in Klasse B (M = 2.42; SD = 0.82). Im Ergebnis sprechen die Befunde für eine hohe Sensibilität der SHA-Schüler:innen gegenüber der lehrerseitigen Gesprächsführung im Unterricht. In der Präsentation werden die Befunde anhand von transkribierten Interaktionssequenzen und Interviewpassagen weiterführend illustriert und diskutiert.

Zum Zusammenhang zwischen bedürfnisunterstützender Unterrichtsgestaltung, Bedürfniserfüllung im Unterricht und dem fachbezogenen Interesse von Grundschüler*innen

Cornelia Reich
Ludwig-Maximilians-Universität München, Germany

Theoretischer Hintergrund und Forschungsstand

Eine an den Schüler*innen ausgerichtete bedürfnisunterstützende Unterrichtsgestaltung zeichnet sich durch Autonomieunterstützung, Struktur sowie Involvement aus (Skinner & Belmont, 1993). Der Basic Needs Theory folgend unterstützt eine solche Unterrichtsgestaltung die Grundbedürfnisse nach Autonomie, Kompetenz und Sozialer Eingebundenheit (Ryan & Deci 2017). Überdies können aus dieser theoretischen Grundlage wichtige Bedingungsfaktoren für fachbezogenes Interesse begründet werden (Krapp 1998; 2007).

Empirische Befunde zeigen, dass Autonomieerfüllung durch Autonomieunterstützung gefördert wird (Noels 2003). Zudem konnten Effekte von Bedürfnisunterstützung auf Interesse nachgewiesen (Hartinger 2006, Oppermann & Lazarides 2021) und positive Zusammenhänge zwischen Erfüllung von Autonomie, Kompetenz und Sozialer Eingebundenheit mit weiteren emotional-motivationalen Faktoren gezeigt werden (Behzadnia 2020, Buzzai & Sorrenti 2021).

Der bisherige Forschungsstand macht verschiedene Desiderata sichtbar: In Bezug auf die überwiegend bei Schüler*innen der Sekundarstufe durchgeführten Untersuchungen (Übersicht: Stroet et al. 2013) ist zunächst offen, inwieweit die gewonnenen Befunde für die Altersgruppe der Grundschüler*innen gelten. Hinsichtlich der bedürfnisunterstützenden Unterrichtsgestaltung wird in bisherigen Studien insbesondere auf die Autonomieunterstützung fokussiert. Wenig erforscht hingegen ist die Unterstützung von Kompetenz und Sozialer Eingebundenheit im Unterricht, speziell in der Primarstufe. Bezogen auf das individuelle Interesse liegen im Grundschulkontext mehrheitlich Befunde zu außerschulischen Interessen vor (Laufs & Kempert 2021, Vogt et al. 2000), jedoch weniger zu fachbezogenen Themeninteressen.

Fragestellung und Hypothesen

Vor dem Hintergrund des Forschungsstands wird die Fragestellung verfolgt, wie die wahrgenommene Bedürfnisunterstützung im Mathematikunterricht, mediiert durch die wahrgenommene Bedürfniserfüllung, mit dem individuellen Mathematikinteresse von Grundschüler*innen zusammenhängt.

H_1*:* Es wird angenommen, dass *Autonomieunterstützung* ($H_{1.1}$), *Struktur* ($H_{1.2}$) und *Involvement* ($H_{1.3}$) positiv mit *Mathematikinteresse* zusammenhängen.

H_2: Der Zusammenhang von *Autonomieunterstützung* mit *Mathematikinteresse* wird durch *Autonomieerfüllung* mediiert.

H_3:llDer Zusammenhang von *Struktur* mit *Mathematikinteresse* wird durch *Kompetenzerfüllung* mediiert.

H_4:lllDer Zusammenhang von *Involvement* mit *Mathematikinteresse* wird durch *Eingebundenheitserfüllung* mediiert.

Methodik

Im Rahmen einer Querschnittserhebung wurde im Schuljahr 2021/2022 eine quantitative Befragung mit Paper-Pencil-Fragebögen von 441 Dritt- und Viertklässler*innen (49% Mädchen, MW_{Alter} = 9 Jahre) durchgeführt.

Zur Erfassung der wahrgenommenen bedürfnisunterstützenden Unterrichtsgestaltung, wahrgenommenen Bedürfniserfüllung sowie des individuellen Mathematikinteresses wurden bereits vorhandene Instrumente mathematikspezifisch und altersadaptiert eingesetzt sowie auf einer fünfstufigen Likert-Skala beantwortet (Admiraal et al. 2019, Thiel et al. 2013, Torsheim et al. 2013, Van der Kaap-Deeder et al. 2020, Willems et al. 2020).

Die Datenanalysen zeigen, dass die Faktorenstrukturen der einzelnen Skalen zufriedenstellende FIT-Indizes ($\chi 2$/df, CFI, RMSEA) aufweisen, weshalb die Analysen mittels Strukturgleichungsmodellierung erfolgen.

Ergebnisse und ihre Bedeutung

Die Analysen zeigen, dass ein positiver Zusammenhang von *Autonomieunterstützung* und *Struktur* mit *Mathematikinteresse* besteht ($H_{1.1}$, $H_{1.2}$), wobei sich *Autonomieerfüllung* (H_2) sowie *Kompetenzerfüllung* (H_3) als signifikante Mediatoren erweisen. Involvement zeigt ebenso hypothesenkonform einen positiven Zusammenhang mit *Mathematikinteresse* ($H_{1.3}$), erwartungswidrig ist jedoch der nicht signifikante Mediatoreffekt von *Eingebundenheitserfüllung* (H_4).

Die ermittelten Befunde geben Hinweise darauf, dass Lehrkräfte im Rahmen der Unterrichtsgestaltung auf die Grundbedürfnisse ihrer Schüler*innen eingehen sollten, um deren Mathematikinteresse zu wecken/aufrecht zu erhalten. Zu diskutieren ist zudem die Bedeutung von *Autonomieerfüllung, Kompetenzerfüllung* sowie des *Involvements* von Lehrkräften – nicht aber von *Eingebundenheitserfüllung* – für das individuelle *Mathematikinteresse*. Weitere Implikationen sowie Limitationen werden im Vortrag diskutiert.

Authentische Geschichte(n)? Wie authentische Lernmaterialien die Involviertheit der Lernenden beeinflusst

Lisa Zachrich[1], Wolfgang Wagner[1], Christiane Bertram[2], Ulrich Trautwein[1]
[1]Universität Tübingen, Germany; [2]Universität Kontsanz, Germany

Der Lebensweltbezug von Lerninhalten ist ein zentrales didaktisches Prinzip und wird in authentischen Lernarrangements, beispielsweise beim Lernen am historischen Ort, im bilingualen Unterricht oder in außerschulischen Forschungslaboren, in besonderem Maße umgesetzt. Authentische Lernarrangements werden mit Blick auf den Lernerfolg als besonders vielversprechend angesehen (Fougt et al., 2019; Rule, 2006). In den letzten Jahren fanden empirische Studien mit unterschiedlichen Fachbezügen allerdings nur selten Hinweise auf die theoretisch angenommenen Erfolge der authentischen Lernarrangements (z.B. Bertram et al., 2017; Itzek-Greulich et al., 2015; Nachtigall & Rummel, 2021). Um zu verstehen, wie es zu den Inkonsistenzen zwischen theoretischen Annahmen und empirischen Ergebnissen kommt, erscheint ein Blick auf die Lernprozesse vielversprechend: Was genau passiert, wenn Lernende mit authentischen Lernarrangements konfrontiert werden? Die vorliegende Studie fußt auf einem Wirkungsmodell von Betz et al. (2016) zum Lernen in authentischen Lernarrangements. Es beschreibt einen Prozess, in dem aus einem Zusammenspiel individueller Lernvoraussetzungen und Eigenschaften der Lernsituation eine Wahrnehmung von Authentizität entsteht. Diese wirkt auf die Motivation und das Lernergebnis. Wir erweiterten dieses Modell um die Komponente „situationale Verarbeitung und Involviertheit“. In Anlehnung an Zachrich et al. (2020) konzeptualisieren wir die situationale Verarbeitung und Involviertheit als kognitive, affektive und physische Verarbeitungsprozesse.

Im Geschichtsunterricht spielt das Lernen mit authentischem Lernmaterial, insbesondere auch das Lernen mit Zeitzeug:innenberichten, eine bedeutsame, aber auch umstrittene Rolle. Denn die authentischen Berichte sollen Lernende für die Geschichte sensibilisieren und gleichzeitig soll der Unterricht sie dazu befähigen, mit diesen reflektiert umzugehen (z.B. Henke-Bockschatz, 2014). Problematisch dabei ist, dass Lernende in der Begegnung mit den Berichten häufig denken, sie lernen *die Wahrheit* und nicht *eine Wahrheit* über die Vergangenheit (z.B. Ballis & Schwendemann, 2021).

Wir nehmen an, dass die Verarbeitung der Lernenden eine zentrale Rolle in der Wirkung von authentischen Lernarrangements auf verschiedene Lernergebnisse spielt. Konkret untersuchten wir, welchen Effekt die Authentizität persönlicher Erfahrungsberichte aus der Vergangenheit auf die situationale Verarbeitung des Berichts, auf das situationale Interesse an Geschichte, die gefühlte Erkenntnis und die kritische Distanz hat.

Zur Beantwortung der Forschungsfrage führten wir eine experimentelle Online-Studie durch, in der wir lediglich das Framing bei der Präsentation von Zeitzeug:innenvideos variierten und so drei Experimentalbedingungen konzipierten: In der Bedingung mit hoher Authentizität (*original*) bekamen Teilnehmende die Information, dass im Videoclip ein:e Zeitzeug:in über persönliche Erfahrungen berichtet (authentische Person, authentischer Inhalt). In der Bedingung mit mittlerer Authentizität (*nachgespielt*) enthielt das Framing die Information, dass im Video ein:e Schauspieler:in den Bericht eines:r Zeitzeug:in nachspielt (nicht-authentische Person, authentischer Inhalt). In der Bedingung mit geringer Authentizität (*rekonstruiert*), gaben wir den Teilnehmenden die Information, dass ein:e

Schauspieler:in einen rekonstruierten Bericht eines:r Zeitzeug:in nachspielt, der laut Historiker:innen so ähnlich stattgefunden haben könnte (nicht-authentische Person, nicht-authentischer Inhalt).

Insgesamt nahmen 453 Personen über die Online-Plattform prolific.org teil, von denen 428 randomisiert wurden. 19.4% beantworteten mindestens einen von zwei Manipulation Checks falsch, weshalb die finale Stichprobe aus N = 336 Teilnehmenden, $n_{original}$ = 120, $n_{nachgespielt}$ = 110, and $n_{rekonstruiert}$ = 106, bestand. Wir nutzten standardisierte Fragebögen (Autor, in Druck) zur Erfassung der situationalen Verarbeitung und führten multiple Gruppenmodelle durch.

Es zeigte sich, dass die Authentizität des Videoclips einen Effekt auf die emotionale Verarbeitung der Lernenden mit dem Material hatte: Lernende in der Original-Bedingung gaben an, stärker emotional involviert zu sein als Lernende in den Bedingungen mit nachgespieltem und rekonstruiertem Bericht. Das situationale Interesse der Teilnehmenden war in der Original-Bedingung statistisch signifikant höher als in den anderen beiden Bedingungen. Für die gefühlte Erkenntnis und die kritische Distanz zum Videoclip fanden wir keine statistisch signifikanten Unterschiede. Herauszustellen ist, dass bereits eine minimale Variation der Videoeinführung substanzielle Effekte auf die emotionale Verarbeitung der Lernenden hat. Das gibt einen Hinweis darauf, wie bedeutsam die emotionalen Verarbeitungsprozesse in dem Lernarrangement mit Zeitzeug:innen sind.

Gemeinschaft im individualisierten Unterricht: Zusammenhänge des Gemeinschaftsgefühls und der sozialen Eingebundenheit mit Merkmalen der individualisierten Unterrichtsgestaltung

Simon Ohl, Hanna Dumont
Universität Potsdam, Germany

Theoretischer Hintergrund

Das Grundbedürfnis nach sozialer Eingebundenheit gilt als fundamentales Motiv menschlichen Handelns und ist auch für das schulische Lernen von Bedeutung (Ryan und Deci 2017). Die soziale Eingebundenheit der Schüler:innen in eine Klassengemeinschaft stellt demnach einen bedeutsamen Faktor für den Schulerfolg dar (Osterman 2000).

In einem individualisierten Unterricht stellt sich die Frage, was das „Gemeinsame“ ist, auf dem eine Klassengemeinschaft beruht und inwiefern alle Schüler:innen Teil dieser Gemeinschaft sind (Bossen und Merl 2021). So haben ethnographische Studien gezeigt, dass Individualisierung teilweise in Form eines „dezentralisierten“ Unterricht umgesetzt wird, in dem statt eines gemeinsamen Lehr-Lernprozesses im Klassenverbund, voneinander abgegrenzte individuelle Lernprozesse nebeneinander stattfinden und der Austausch zwischen Schüler:innen stark eingeschränkt ist (Rabenstein et al. 2018).

Werden individuelle Lernprozesse hingegen aufeinander bezogen, bietet sich die Gelegenheit, dass Lerngegenstände aus unterschiedlichen Perspektiven beleuchtet werden und die Schüler:innen sich gegenseitig beim Lernen unterstützen (Leuders und Prediger 2017). Damit dies gelingt, sollten die sozialen Beziehungen zwischen den Schüler:innen so gestalten werden, dass eine kooperative und wertschätzende Lerngemeinschaft entsteht, die alle Schüler:innen inkludiert und zum Lernen motiviert (Farmer et al. 2019).

Fragestellung

Zentrale Fragestellung des vorliegenden Beitrags ist, ob eine Individualisierung des Unterrichts mit weniger Gemeinschaft im Unterricht einhergeht. Dazu werden zwei Konstrukte herangezogen, die Gemeinschaft im Schul- und Unterrichtskontext auf unterschiedliche Weise abbilden: das Gemeinschaftsgefühl der Lerngruppe (GG) und die situative soziale Eingebundenheit der Schüler:innen (SSE). Es wird angenommen, dass beide Konstrukte positiv miteinander korrelieren, aber unterschiedliche Aspekte von Gemeinschaft erfassen. Außerdem wird angenommen, dass eine hoch ausgeprägte Individualisierung des Unterrichts mit einer niedrigeren Ausprägung der beiden Konstrukte einhergeht.

Methode

Zur Beantwortung der Fragestellung wurden Daten aus einer Studie herangezogen, die an neun Grundschulen durchgeführt wurde, die mit dem Deutschen Schulpreis ausgezeichnet wurden und sich durch eine individualisierte Unterrichtsgestaltung auszeichnen. Die Stichprobe umfasst N=563 Schüler:innen der dritten und vierten Jahrgangsstufen aus 49 Lerngruppen. Für das GG wurde eine, von den Autor:innen entwickelte, Fragebogenskala zur Erfassung des Gemeinschaftsgefühls verwendet (10 Items, Bsp.: „Bei uns im Unterricht fühlen wir uns miteinander verbunden.“, Cronbachs’ α=0.87, ICC1=0.16), für das SSE wurde anhand der Experience-Sampling-Methode in 23 der 49 Lerngruppen über eine Woche hinweg mehrfach pro Tag erfasst, ob die Schüler:innen sich im Unterricht sozial eingebunden fühlten (Single-Item: „Ich habe mich gerade wohl mit meinen Mitschülerinnen und Mitschülern gefühlt“, ICC1=0.03). Zunächst wurde die Korrelation zwischen GG und SSE auf Individual- und Gruppenebene berechnet. Anschließend wurden Mehrebenen-Regressionsmodelle auf Individual- und Gruppenebene mit GG bzw. SSE als abhängige Variablen und Skalen zur individualisierten Unterrichtsgestaltung (Differenzierung, α=0.82, ICC1=0.61; Autonome Aufgabenbearbeitung α=0.77, ICC1=0.18; Selbstgesteuertes Lernen, α=0.83, ICC1=0.32) als Prädiktoren durchgeführt. Alle Variablen wurden anhand der Gesamtvarianz z-standardisiert.

Ergebnisse & Diskussion

Es zeigte sich eine positive Korrelation zwischen GG und SSE auf Individualebene (r=0.26, SE=0.07) und Gruppenebene (r=0.62, SE=0.24). Wie angenommen fühlten sich die Schüler:innen in Lerngruppen, in denen das Gemeinschaftsgefühl hoch ausgeprägt war, auch stärker sozial eingebunden. Die schwache bzw. moderate Stärke der Korrelation auf Individual- bzw. Gruppenebene deutet darauf hin, dass von GG und SSE unterschiedliche Aspekte von Gemeinschaft erfasst werden.

Keine der Skalen zur individualisierten Unterrichtsgestaltung hing negativ mit GG oder SSE zusammen. Im Gegenteil zeigten sich positive Zusammenhänge zwischen GG und autonomer Aufgabenbearbeitung auf Individualebene (β=0.21, SE=0.05, R^2_{Level1}=0.04), nicht jedoch auf Gruppenebene. Die SSE hing positiv mit autonomer Aufgabenbearbeitung (β=0.17, SE=0.08, R^2_{Level2}=0.31) sowie mit Differenzierung (β=0.10, SE=0.05, R^2_{Level2}=0.30) jeweils auf Gruppenebene, nicht jedoch auf Individualebene zusammen. Die Annahme, dass eine Individualisierung des Unterrichts mit weniger Gemeinschaft im Unterricht einhergeht, konnte demnach empirisch nicht bestätigt werden.

Gemeinschaft im individualisierten Unterricht ist ein viel diskutiertes Thema, zu dem es bisher kaum empirische Studien gibt. Die vorliegende Studie liefert erste Erkenntnisse die darauf hindeuten, dass Gemeinschaft auch in einem individualisierten Unterricht möglich ist.

G-07: Lehrmaterial und Training

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S05 T00 B59
Chair der Sitzung: Stefan Christian Ferner

Technologiebasierte Förderung der analytischen Problemlösekompetenz (aPLK) auf Basis des Cognitive Apprenticeship Ansatzes und durch das adaptive tutorielle Feedback

Stefan Christian Ferner, Christoph Brauns, Pia Schäfer, Felix Walker
Universität Hamburg, Deutschland

Theoretischer Hintergrund

Im Rahmen der Teilstudie 2 des Forschungsprojekts TechKom (Technologiebasierte Kompetenzmessung und -förderung in der Erstausbildung in der Elektro- und Metalltechnik; Förderkennzeichen: 21AP011), wird die Wirksamkeit einer Intervention zur Förderung der analytischen Problemlösekompetenz (aPLK) untersucht (Walker et al. 2016). Die Zielgruppe sind Auszubildende in den Berufen Elektroniker*in für Automatisierungstechnik und Mechatroniker*in.

Im Allgemeinen wird unter der PLK im technischen Bereich die Fähigkeit verstanden, das Fachwissen in problemhaltigen und variierenden Anforderungssituationen anzuwenden (Walker et al., 2015). Somit äußert sich das Problemlösen als Prozess des Erkennens und Deutens von Zuständen an elektrotechnischen Anlagen (u.a. Harms, 2007).

Die Notwendigkeit zur Förderung der aPLK von Auszubildenden im technisch-gewerblichen Bereich resultiert aus verschiedenen Gründen. Zum einen äußert sich in Folge der Digitalisierung und Automatisierung der industriellen Fertigungsprozesse eine gestiegene Anforderung zukünftiger Instandhalter*innen (Gutsche & Voigt, 2018). Zum anderen zeigte sich am Ende der dualen Erstausbildung von Elektroniker*innen für Automatisierungstechnik eine Diskrepanz zwischen curricularen Vorgaben und der tatsächlich vorhandenen Kompetenzausprägungen (Walker et al., 2016). Daraus resultiert ein Förderbedarf, dem in der Teilstudie 2 des Forschungsprojekts TechKom entgegnet werden soll.

Ziel der Studie ist die Untersuchung des Einflusses adaptiv-tutoriellen Feedbacks und der kognitiven Modellierung auf die Entwicklung der aPLK. Unter dem adaptiv-tutoriellen Feedback werden Informationen und Hilfestellungen verstanden, die sich dem Lernbedarf situativ anpassen und nur das bereitstellen, was tatsächlich benötigt wird (Narciss, 2006).

Fragestellung

Aus der aktuellen Forschungslage stellten sich zwei Forschungsfragen, die sich in den nachfolgenden Hypothesen aufstellen ließen:

H1: Auszubildende, welche in der Phase „modeling“ nur auf das Video zur kognitiven Modellierung zurückgreifen können, erwerben weniger aPLK als Auszubildende, die während der Bearbeitung von Problemen ATF in der Computersimulation erhalten.

H2: Auszubildende, welche im AIST-R Personencode eine technische Orientierung aufweisen, diagnostizieren Fehlerfälle erfolgreicher, als Auszubildende, welche keine technische Orientierung erzielen.

Methode

Zur Erfassung der aPLK wurde eine Automatisierungsanlage digital simuliert. Die Auszubildenden nutzen die Simulation einer industrienahen Automatisierungsanlage, kurz SINA (Walker et al. 2016), um eine strategiegeleitete Fehlersuche durchzuführen (Benda 2008; Konradt 1995). Zu diesem Zweck wurden verschiedene Fehlerfälle entwickelt und in die Simulation SINA integriert. Diese müssen von den Lernenden identifiziert und gelöst werden. Bei den Fehlerfällen handelt es sich um Defekte an der Hardware der Anlage, sowie um Fehler in der Programmierung des SPS-Programms.

Die Entwicklung der Programmfehler basiert auf empirisch geprüften Fehlern von Walker et al. (2016). Zur Förderung der aPLK wurden Lösungsvideos zur kognitiven Modellierung entwickelt (z.B. Schaper, Hochholdinger & Sonntag, 2003) und allen Auszubildenden präsentiert. Die Lösungsvideos basieren auf dem Cognitive Apprenticeship Ansatz (Collins, Brown & Newman 1989) und beinhalten Strategien zur Lösung fachspezifischer Probleme.

In einem Experimental (EG)-Kontrollgruppen (KG) - Design wird die Experimentalgruppe zusätzlich durch adaptiv-tutorielles Feedback unterstützt (z.B. Marschner 2011; Narciss 2006), dessen Wirksamkeit u.a. von Schaper et al. (2003) nachgewiesen wurde. Das ATF wurde in die Simulation integriert und in Form eines Chats zwischen den Probanden und einem*r Instandhalter*in realisiert.

Sowohl die Förderung der aPLK als auch die Datenerhebung fanden auf einer speziell entwickelten digitalen Lernplattform (digilearn.tech) statt.

Ergebnisse

Anfang 2022 erfolgte die Pilotierung mit 102 Auszubildenden (54 Mechatroniker*innen, 48 Elektroniker*innen für Automatisierungstechnik).

Die ersten Ergebnisse zeigen in beiden Gruppen einen signifikanten positiven Einfluss der Intervention in beiden Gruppen (Eta-Quadrat = .224, $p < .001$). Im Rahmen der Pilotierung konnte jedoch kein signifikanter Einfluss der Gruppenzugehörigkeit auf die analytische Problemlösekompetenz nachgewiesen werden (H1). Die Kombination aus Videos zur kognitiven Modellierung und adaptiv tutoriellem Feedback (EG) weist demnach in der ersten Untersuchung keinen höheren Einfluss auf die analytische Problemlösekompetenz auf als der ausschließliche Konsum von Videos zur kognitiven Modellierung (KG).

Die Interessensdimensionen (H2) auf die erfolgreiche Fehlerdiagnose nahm statistisch keinen signifikanten Einfluss ein.

Diese Ergebnisse werden in der für Oktober 2022 geplanten Haupterhebung überprüft.

Risiken evidenzbasiert einschätzen lernen - Ein systematischer Vergleich von Trainingsmethoden zum Bayesianischen Denken im Projekt TrainBayes

Nicole Steib[1], Theresa Büchter[2], Karin Binder[3], Katharina Böcherer-Linder[4], Andreas Eichler[2], Stefan Krauss[1], Markus Vogel[5]

[1]Universität Regensburg, Deutschland; [2]Universität Kassel, Deutschland; [3]Ludwig-Maximilians-Universität München, Deutschland; [4]Universität Freiburg, Deutschland; [5]Pädagogische Hochschule Heidelberg, Deutschland

Theoretischer Hintergrund

Bayesianisches Denken ist oft mit Missverständnissen verbunden, die regelmäßig zum Beispiel in der Medizin zu falschen Diagnosen oder zu fehlerhaften Gerichtsentscheidungen in der Rechtsprechung

führen. Bei einer typischen Bayesianischen Aufgabe wird ein Risiko (z.B. positiv prädiktiver Wert) auf Basis von drei statistischen Informationen zur Evidenz (z.B. Basisrate, Richtig-Positiv-Rate und Falsch-Positiv-Rate) berechnet. Die Fähigkeit eine solche Wahrscheinlichkeit berechnen zu können wird unter dem Begriff „Performanz“ verstanden (Büchter et al., 2022). Eine Metaanalyse (McDowell & Jacobs, 2017) zeigt, dass lediglich 4% der Personen eine solche Aufgabe im Wahrscheinlichkeitsformat lösen können.

Zwei Strategien haben sich bei solchen Bayesianischen Aufgaben als hilfreich erwiesen: 1) Die Präsentation der statistischen Informationen in Form von natürlichen Häufigkeiten anstatt von Wahrscheinlichkeiten (Gigerenzer & Hoffrage, 1995) und 2) Visualisierungen, z.B. Doppelbäume oder Einheitsquadrate (Binder, Krauss & Wiesner, 2020; Eichler, Böcherer-Linder & Vogel, 2020).

Obwohl Bayesianisches Denken in vielen Professionen relevant ist und Fehlurteile fatale Konsequenzen haben können, gibt es bisher kaum empirische Belege für adressatengerechte Trainings, die die beiden bislang erfolgreichsten Strategien natürliche Häufigkeiten und Visualisierungen kombinieren und systematisch vergleichen.

Im Vortrag wird eine Studie vorgestellt, in der verschiedene Trainings empirisch verglichen werden.

Methode

Im Projekt TrainBayes (http://www.bayesianreasoning.de/br_trainbayes.html) werden zwei sog. „Idealtrainings“ (mit Visualisierung Doppelbaum oder Einheitsquadrat *und* natürlichen Häufigkeiten) und zwei Alternativtrainings (nur mit natürlichen Häufigkeiten oder einer curricularen Umsetzung) in einem Prä-Post-Follow-up-Test-Design bei Medizinstudierenden (N=260) und Jurastudierenden (N=255) untersucht. In den Trainings wird zunächst gelernt, wie der positiv prädiktive Wert berechnet werden kann. Anschließend wird in einer Übungsphase mit individuellem Feedback das Gelernte an einem authentischen Kontext vertieft. Die Trainingsmaterialien sind in Büchter et al. (2022) detailliert beschrieben.

Ergebnisse und ihre Bedeutung

Der Anteil korrekter Aufgaben verändert sich in den vier Trainingsgruppen von Prätest (Doppelbaum=12,6%, Einheitsquadrat=18,7%, Häufigkeiten=15,5%, curriculare Umsetzung=13,8%) zu Posttest (Doppelbaum=70,5%, Einheitsquadrat=55,7%, Häufigkeiten=55,1%, curriculare Umsetzung=56,5%) bzw. zu Follow-up-Test (Doppelbaum=55,8%, Einheitsquadrat=35,9%, Häufigkeiten=33,8%, curriculare Umsetzung=36,2%) deutlich, während der Anteil korrekter Aufgaben in der Kontrollgruppe deskriptiv geringer zwischen Prätest (12,9%), Posttest (24,5%) und Follow-up-Test (24,1%) variiert. Vor allem das Einheitsquadrattraining, Häufigkeitstraining und das curriculare Training bewirken ähnliche Lernzuwächse von Prätest zu Posttest (die Follow-Up-Test-Performanzen sinken erwartungsgemäß wieder). Beim Doppelbaumtraining wurde sogar im Follow-up-Test noch über die Hälfte der Aufgaben korrekt beantwortet.

Um den Effekt der Gruppenzugehörigkeit (Kontrollgruppe, Doppelbaumtraining, Einheitsquadrattraining, Häufigkeitstraining oder Curriculares Training) im Prä-Post-Follow-up-Design statistisch zu vergleichen, wurde ein gemischtes lineares Modell mit Maximum-Likelihood-Methode

geschätzt. In diesem Modell wurde das Häufigkeitstraining mit der Performanz im Prätest als Referenzgruppe verwendet und die Faktoren „Posttest“, „Follow-up-Test“ sowie „Kontrollgruppe“, „Doppelbaum“, „Einheitsquadrat“ und „Curriculare Umsetzung“, mittels Dummy-Kodierung als Prädiktoren einbezogen.

Die (nicht-standardisierten) Regressionsgewichte für den Faktor Messzeitpunkt (Prä-Post-Vergleich: b_1=0.396, p<0.001; Prä-Follow-up-Vergleich: b_2=0.183, p<0.001) sind in der Häufigkeitsbedingung signifikant. Die Interaktion für den Vergleich von Zuwächsen in Häufigkeits- und Kontrollbedingung ist für den Prä-Post-Vergleich (b_3=– 0.28, p<0.001), aber nicht für den Prä-Follow-up-Vergleich (b_4=–0.07, p=0.132) signifikant. Die Performanz in der Gruppe mit Häufigkeitstraining wird also kurzfristig, aber nicht langfristig gegenüber der Kontrollgruppe verbessert. Die Regressionsgewichte für die Interaktion zum Vergleich von Zuwächsen in Häufigkeits- und Einheitsquadratbedingung bzw. Häufigkeitsbedingung und curricularer Bedingung werden alle nicht signifikant, also weicht der Lernzuwachs mit keinem dieser Trainings inferenzstatistisch von dem mit dem Häufigkeitstraining ab. Allerdings wird die Interaktion zum Vergleich von Zuwächsen in Häufigkeits- und Doppelbaumbedingung sowohl für den Prä-Post-Vergleich (b_5=0.183, p<0.001) als auch den Prä-Follow-up-Vergleich (b_6=0.249, p<0.001) signifikant. Mit dem Doppelbaumtraining wird also kurz- und langfristig signifikant besser gelernt als mit dem Häufigkeitstraining.

Die Ergebnisse zeigen: Medizin- und Jurastudierende können mit der Trainingsstrategie „Kombination aus Doppelbaum und natürlichen Häufigkeiten“ langfristige Lernerfolge im für sie relevanten Bayesianischen Denken erzielen. Erstaunlich ist auch der Dropout von 0 Studierenden zum Follow-up-Test, der als Beleg für die empfundene Berufsrelevanz interpretiert werden kann.

Video oder Text - Wie wird die Wahl des Lernmaterials beeinflusst?

Sabine Schlag, Sabine Glock
Bergische Universität Wuppertal, Germany

Im Zug der stärkeren Digitalisierung von Unterricht, auch bedingt durch die Maßnahmen der Corona-Pandemie, kommen mehr Videos und andere digitale Materialien zum Einsatz (Wößmann er al. 2020). Aber wie entscheiden Lehrkräfte, welchen Schüler:innen sie welche Materialien zur Verfügung stellen?

Obwohl Lehrkräfte täglich vor der Wahl stehen, Lernmaterial auszuwählen, gibt es keine uns bekannte Forschung dazu, wie diese Wahl beeinflusst wird. Wir gehen davon aus, dass Einstellungen zur Repräsentationsform (Text und Video) und Schülermerkmale (hohe vs. niedrige Leistungsfähigkeit und mit vs. ohne Migrationshintergrund) entscheidend sind.

Um die Einstellungen erfassen zu können, wird auf den Ansatz der epistemologischen Überzeugungen zurückgegriffen. Epistemologische Überzeugungen sind dabei definiert als Einstellung zur Natur von Wissen (Stahl & Bromme, 2007, S. 773). Zusammenhänge zwischen epistemologischen Überzeugungen und dem Lernerfolg sowie anderen lernrelevanten Faktoren sind belegt (zusammenfassend siehe Schommer, 1994). Es wird darüber hinaus auch ein Zusammenhang zwischen den epistemologischen Überzeugungen von Lehr*innen und Schüler*inne sowie dem Unterrichtshandeln der Lehrkräfte angenommen (z.B. Schröder, 2010). Hierzu liegen aber kaum

Untersuchungen vor. Neben den epistemologische Überzeugungen zum Material könnten auch die Überzeugungen der (angehenden) Lehrkräfte zum Umgang der Schüler:innen mit dem Material entscheidend sein. Angelehnt an die Studie von Salomon (1984) gehen wir davon aus, dass (angehende) Lehrkräfte die mentale Anstrengungsbereitschaft von Schüler:innen für Text und Video unterschiedlich einschätzen. Salomon (1984) konnte zeigen, dass die mentale Anstrengungsbereitschaft der Schüler:innen bei Videos geringer war als bei Texten und diese entsprechend auch geringere Lernleistungen zeigten. Es wird deshalb davon ausgegangen, dass auch (angehende) Lehrkräfte von solchen Überzeugungen bei der Wahl von Text- und Videomaterial beeinflusst werden. Nicht nur die Einstellungen zum Lernen mit Text und Video sollten einen Einfluss auf die Auswahl von Lernmaterial haben, sondern auch Stereotype gegenüber Schüler*innen spielen bei der Entscheidungsfindung eine Rolle (Bonefeld & Dickhäuser, 2018). Glock und Krolak-Schwerdt (2013) konnten z.B. zeigen, dass die Bewertungen von angehenden Lehrkräften vom vorhanden sein eines Migrationshintergrunds der Schüler:innen beeinflusst werden. Darüber hinaus könnte auch die Leistungsfähigkeit der Schüler:innen die Wahl des Lernmaterials beeinflussen.

In unserer Studie haben wir untersucht, inwieweit die Wahl von Text- und Videomaterial der angehenden Lehrkräfte dadurch beeinflusst wird, ob der Schüler mit hoher oder niedriger Leistungsfähigkeit sowie mit oder ohne Migrationshintergrund beschrieben wird. Es wurden mit den übersetzten Items von Stahl und Bromme (2007) die epistemologischen Überzeugungen zu Wissen aus Text und Video erfasst (z.B. Wissen aus Videos ist einfach vs. komplex; 7-stufige Likert-Skala, 24 Item-Paare). Außerdem wurden die Fragen zur Attribution sowie der mentalen Anstrengungsbereitschaft nach Salomon (1984), angepasst an die Zielgruppe, abgefragt. Die mentale Anstrengungsbereitschaft wurde nach der Präsentation des inhaltlich äquivalenten Text- und Videomaterials zum Thema Wasserkreislauf bezogen auf präsentierte Material erfasst, z.B. „Wie viel Anstrengung bringen Schüler:innen auf um das Video zu verstehen?“. Es folgt die Beschreibung des Schülers, wobei die angehenden Lehrkräfte eine von vier Fallbeschreibungen erhielten, die darin variierten, ob der Schüler Felix oder Murat hieß (Kleen & Glock, 2020) und ob dieser als leistungsstark oder leistungsschwach beschrieben wurde. Die Proband:in (N = 167) bewerteten abschließend, welches Material sie dem Schüler geben würden und wurden nach der Attribution für die Wahl gefragt.

Die Ergebnisse zeigen, dass die Wahl des Lernmaterials primär von der beschriebenen Leistungsfähigkeit der Schüler abhängig zu sein scheint. Dabei wurde für leistungsschwach beschriebene Schüler mit signifikant höherer Wahrscheinlichkeit das Video ausgewählt (90,8% der Fälle) und für leistungsstark beschriebene Schüler signifikant häufiger der Text (76,6% der Fälle). Weder der Migrationshintergrund noch die anderen Variablen haben einen signifikanten Einfluss auf die Auswahl.

Diese Ergebnisse sollen im Vortrag diskutiert werden. Die Befunde deuten darauf hin, dass angehende Lehrkräfte dem Titel von Salomon (1984): Television is "easy" and print is "tough" immer noch bei ihrer Wahl des Materials folgen.

Julia Harenz[1], Markus Vogel[2], Kristin Altmeyer[1], Roland Brünken[1], Sarah Malone[1]
[1]Universität des Saarlandes, Germany; [2]Pädagogische Hochschule Heidelberg

Lernen mit multiplen externalen Repräsentationen (MER) findet von der Grundschule bis zur Hochschule in unterschiedlichen Bereichen statt. Unter MER versteht man die Kombination verschiedener Repräsentationen im Lern- oder Testmaterial (Schüler, 2020). Multimedia-Lernen bezeichnet laut Mayer (2014) eine spezifische Kombination von Repräsentationen, nämlich Text und Bild, die, wenn sie beim Lernen gemeinsam präsentiert werden, im Vergleich zum Lernen mit nur Text vorteilhafter sind (Multimediaeffekt; Mayer, 1997). Im Lernkontext von MINT-Fächern spielen neben Text und Bild aber auch Formeln eine bedeutsame Rolle. Bisher nur selten wurde das Lernen mit Formeln als eine Repräsentationsform in einem multirepräsentationalen Lernkontext untersucht. 2018 begannen Ott et al. damit, den Einfluss von verschiedenen Repräsentationskombinationen (Text, Formel und Graphik) im Kontext der Aussagenlogik bei Erwachsenen zu untersuchen. Sie klassifizierten dazu MER in heterogene MER und homogene MER. Heterogene MER bestehen aus der Kombination verschiedener Repräsentationstypen – also analogen und symbolischen Repräsentationen, wie z.B. Text-Bild-Kombinationen (Malone et al., 2020). Die Kombination von ausschließlich analogen oder ausschließlich symbolischen Repräsentationen bezeichnen die Autoren als homogene MER. Ein zentrales Ergebnis dieser Arbeiten ist, dass die Kombination von Text und Formel (homogen-symbolische MER) Einzelrepräsentationen überlegen ist und zu ähnlich guten Ergebnissen führt, wie die Kombination von Text und Graphik (heterogene MER). Malone et al. (2020) fand ähnliche Ergebnisse für das Lösen von Gleichungssystemen bei Erwachsenen. Der Ansatzpunkt der vorliegenden Studie ist zu untersuchen, wie Schüler*innen der 4. bis 6. Klasse von homogenen und heterogenen MER profitieren.

In einer Pilotstudie mit 30 Kindern wurde ein eigens entwickeltes Lernprogramm zu aussagenlogischen Verknüpfungen mit entsprechendem Testprogramm am PC vorgetestet. Im Lernprogramm lernen die Teilnehmenden aussagenlogische Verknüpfungen mithilfe von Emojis, die gewisse Eigenschaften haben können oder nicht, kennen und bearbeiten eigenständig nach jeder Verknüpfung zwei Übungsaufgaben mit unmittelbarem Feedback. Die Auswertung erfolgt mit Teilpunkten, sodass für jede Aufgabe der korrekt gelöste Anteil berechnet wurde (M=.79; SD=.077). Im Testprogramm mussten die Teilnehmenden entscheiden, ob Kreise angemalt sind oder nicht (M=.82; SD=.26). Die Ergebnisse beider Programme lassen darauf schließen, dass die Teilnehmenden gut mit den Aufgaben zurechtkamen. Von den 25 entwickelten Testitems benötigten wir für die Hauptstudie nur 12. Da es keine signifikanten Unterschiede in der Itemschwierigkeit gab, wurden diese für die Hauptstudie aufgrund einer inhaltlichen Breite ausgewählt.

Analog zur Studie von Ott et al. (2018) wurde die Hauptstudie – bestehend aus zwei Testterminen – als Between-Subject-Design mit sechs Bedingungen (1: Text, 2: Formel, 3: Text + Formel, 4: Text + Graphik, 5: Formel + Graphik, 6: Text + Formel + Graphik) konzipiert. Die Stichprobe besteht aus 288 Kindern der Klassenstufe 4 bis 6 und wird in Gruppentestungen durchgeführt. Beim ersten Termin bearbeiten die Schüler*innen sowohl nach dem Lernprogramm als auch nach dem Testprogramm einen differenzierten Cognitive-Load-Fragebogen (modifiziert nach Klepsch et al., 2017) bearbeitet. An das Testprogramm schließt sich zusätzlich ein Fragebogen an, welcher die persönlich

empfundene Nützlichkeit der zur Verfügung gestellten Repräsentation(en) abfragt. Beim zweiten Termin bearbeitet die Stichprobe zwei Untertests des BEFKI 5-6 zum schlussfolgernden Denken (verbaler und figuraler Untertest) (Schroeders et al., 2020), den BIS HB und zum Mathematikinteresse, sowie demographische Angaben.

Die Datenerhebung läuft zurzeit noch. Die Studie untersucht, welche Auswirkungen verschiedene Typen von MER, insbesondere unter Einbezug von Formeln, auf die Lerngerbnisse und Kognitive Belastung von Schüler*innen der Klassenstufen 4 bis 6 haben und welche individuellen Voraussetzungen diesen Effekt möglicherweise moderieren. Diese Studie dient als Grundlage, um die Forschung im Bereich Lernen mit MER unter Einbezug von Formeln zu vertiefen.

Gleichzeitig hat diese Studie eine unmittelbare praktische Bedeutung, da ein erstes, wirksames digitales Lernprogramm für zur Vermittlung aussagenlogischer Konzepte für diese Altersstufe entwickelt wurde, das z.B. im Rahmen des kürzlich eingeführten verpflichtenden Informatikunterrichts eingesetzt werden könnte.

G-08: Feedback in der Hochschule

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S05 T00 B71
Chair der Sitzung: Christian L. Paulus

KI im Einsatz? Eine experimentelle Studie zum Feedback in der Hochschullehre

Christian L. Paulus, Benjamin Brummernhenrich, Regina Jucks
Westfälische Wilhelms Universität Münster, Germany

Theoretischer Hintergrund

Feedback liefert Lerner*innen wichtige Hinweise zum eigenen Lernerfolg und beeinflusst zudem auf motivationaler Ebene die weitere Auseinandersetzung mit dem Lerngegenstand. In der Hochschullehre findet Feedback in vielfältiger Weise statt, z.B. als (unstrukturierte) Peer-to-Peer oder schriftliche Rückmeldung. Auch der Individualisierungsgrad von Feedback variiert.

Mit Methoden der Künstlichen Intelligenz können z.B. durch Chatbots Feedbacksituationen gestaltet werden und eine größere Zahl an Feedback bereitgestellt werden, als dies bei einer individualisierten Rückmeldung durch Lehrende (oder Tutor*innen) möglich wäre.

In vorhergehenden Studien wurde untersucht, wie der Kommunikationsstil von Spoken Dialogue Systems (SDS) sich auf deren Wahrnehmung durch Menschen auswirkt. Menschliche Kommunikationspartner wurden im Regelfall bevorzugt. Ein menschlicher Kommunikationspartner wurde mehr gemocht und die Kommunikation wurde als weniger kognitiv beanspruchend bewertet. Allerdings führte ein elaborierter Kommunikationsstil dazu, dass auch die SDS als angenehme Kommunikationspartner beurteilt wurden (Linnemann & Jucks, 2016; Jucks et al., 2018). Insbesondere die Berücksichtigung von Politeness (Brown & Levinson, 1987), also die Vermeidung von beispielsweise autonomie-einschränkenden Handlungsaufforderungen, führte zu höheren Bewertungen der Computersysteme. Bei einer Feedback-KI stellt sich also die Frage, ob auch ihre Informationen von den Studierenden als präziser oder höherwertiger bewertet würden und inwiefern sich der Kommunikationsstil auf die Wahrnehmung auswirkt.

Analog zu diesen Studien greifen wir auf das das Muenster Epistemic Trustworthiness Inventory (METI) als Messinstrument zurück (Hendriks et al., 2015). METI baut auf dem ABI-Modell von Mayer, Davis und Schoorman (1995) auf, welche Ability, Benevolence und Integrity in ihrem Theorieansatz als fundamentale Bausteine für Vertrauen etablierten. METI eröffnet also drei (Expertise, Wohlwollen und Integrität) unterschiedliche Facetten von Vertrauenswürdigkeit bei Fach- und Kompetenzpersonen.

Als weiteres Instrument haben wir den Big-Two-Ansatz von Agency und Communion in die Untersuchung mit einfließen lassen (Abele & Bruckmüller, 2013). Dieser eignet sich, da Communion eine Dimension darstellt, welche soziale Nähe zwischen Beobachter und Beobachtetem abbildet. Agency hingegen drückt beispielsweise die wahrgenommene Zielstrebigkeit und Autonomie aus, zeigt dabei hohe Korrelationswerte mit dem wahrgenommenen sozialen Status einer Person. Diese beiden Dimensionen erweitern in der Studie das Blickfeld dafür, wie eine KI als (unterstützende) Lehrkraft an der Hochschule wahrgenommen wird.

Forschungsfrage

In einer Studie mit 2x2 between-subject design sind wir mit dem Fokus der Wahrnehmung einer KI als Feedbackgeber folgender Forschungsfrage nachgegangen: Wie wird eine KI, im Vergleich zu einem normalen Lehrenden, als Feedbackgeber in einem Chat-Kontext wahrgenommen? Und inwiefern wirkt sich die Berücksichtigung von Politeness-Kommunikationsstrategien auf die Wahrnehmung der Feedbackgeber aus?

Methode

Wir haben gegen Ende 2021 eine Online-Studie mit 284 Studierenden aus NRW durchgeführt: Die Teilnehmer*innen bekamen einen fiktiven Chat-Dialog zwischen einem Studenten und einem Feedbackgeber vorgelegt, bei dem eine schriftliche Einreichung evaluiert wurde. Bei dem Dialog wurde einerseits die Autorenschaft (Lehrender vs. KI-System), andererseits die Ausgestaltung des Feedbacks (wertschätzend/höflich vs. direkt/ungeschönt) manipuliert. Im Anschluss füllten die Teilnehmer*innen einen Fragebogen dazu aus, wie sie die Vertrauenswürdigkeit des Feedbackgebers wahrgenommen haben und welche Adjektive (aus dem Agency & Communion-Paar) sie für den Feedbackgeber als zutreffend einordnen würden.

Ergebnisse

In der Auswertung zeigten sich vier Haupteffekte durch die Manipulationen: Dem KI-System wurde weniger Wohlwollen zugesprochen als dem Lehrenden. Darüber hinaus sorgte die höflich, wertschätzende Formulierungsversion des Feedbacks dafür, dass ein Feedbackgeber (KI und Lehrender) als wohlwollender wahrgenommen wird. Gleiches gilt für die wahrgenommene Communion, welche durch höfliches Feedback erhöht beurteilt wird. Ein direktes, ungeschöntes Feedback führt dazu, dass bei dem Feedbackgeber eine höhere Agency wahrgenommen wird.

Im Hinblick auf die Forschungsfrage, inwiefern eine KI als Feedbackgeber (anders) gesehen wird, zeigt sich ein signifikanter Effekt bezüglich des Wohlwollens. Allerdings deuten auch die Ergebnisse bei Expertise und Agency darauf hin, dass sich hier kleine Effekte zu Ungunsten der KI finden lassen könnten.

Die Ergebnisse werden mit Blick auf den Einsatz von KI in der Hochschullehre diskutiert.

Angemessenheit von Peer-Feedback: Welche Rolle spielen individuelle Merkmale der Lernenden?

Martin Greisel, Julia Hornstein, Ingo Kollar
Universität Augsburg, Deutschland

Peer-Feedback ist eine ökonomische Lösung, um insbesondere in der Hochschulbildung lernwirksames Feedback (Hattie & Clarke, 2018) realisieren zu können. Unklar ist aber, inwiefern individuelle Charakteristika beeinflussen, ob ein Feedback als angemessen wahrgenommen wird. Dies ist aber wichtig für den Lernerfolg, denn je angemessener Feedback erscheint, desto tiefer und damit nutzbringender verarbeiten Feedbackempfänger:innen es (Berndt et al., 2022; Jurkowski, 2018). Diese Angemessenheitseinschätzung sollte einerseits von der tatsächlichen Qualität des Feedbacks abhängen, welches wiederum von den Eigenschaften der Feedback-Geber:innen mitbestimmt wird. Dies sind insbesondere die empfundene Verantwortung, gutes Feedback zu geben (Strijbos et al., 2021), die Zuversicht, gutes Feedback geben zu können, und die generelle

Wertschätzung von Peer-Feedback als Instruktionsmethode und der dafür notwendigen Fähigkeiten (Huisman et al., 2020). Andererseits sollten in die Angemessenheitseinschätzung Merkmale der Feedback-Empfänger:innen, wie deren Empfänglichkeit für auch kritisches Feedback (Strijbos et al., 2021) oder der Glaube, dass Peers in der Lage sind, hochwertiges Feedback zu geben (Huisman et al., 2020), einfließen.

Die Studie untersucht daher, wie feedbackbezogene Lernendenmerkmale mit der Wahrnehmung der Feedbackqualität zusammenhängen.

Im Rahmen einer Lehrveranstaltung mit pädagogisch-psychologischen Inhalten haben 254 Lehramtsstudierende (78% weiblich; *M* = 22,56 Jahre alt [*SD* = 4,30]; im *M* = 4,61 [*SD* = 1,20] Semester) freiwillig an einer Online-Längsschnittstudie teilgenommen. In einem Vortest wurden die genannten Lernendenmerkmale erfasst. Anschließend bearbeiteten die Studierenden individuell eine Vignette einer Unterrichtsstunde, in der die Lehrkraft mit mehreren Problemen konfrontiert war. Die Aufgabe der Studierenden war, diese Probleme aus pädagogisch-psychologischer Sicht zu identifizieren, mithilfe von pädagogisch-psychologischen Theorien zu analysieren, Ziele für die Lehrkraft abzuleiten und Verbesserungsvorschläge zu entwickeln. In der Folgewoche wurden diese Analysen dann zufällig je zwei Peers zugeordnet, die dazu schriftliches Feedback gegeben haben. Dieses Feedback wurde dann wiederum eine Woche später von den Feedback-Empfänger:innen gelesen und bewertet. Zuletzt haben diese auf Basis des Feedbacks ihre ursprüngliche Fallanalyse überarbeitet.

Mit der Peer-Feedback Orientation Scale (Kasch et al., 2022) wurde gemessen, wie verantwortlich man sich fühlt, hilfreiches Peer-Feedback zu geben (Accountability; ω_h =.70), bzw. wie offen man sich einschätzt, mit unterschiedlichen Standpunkten oder positivem wie negativem Feedback gleichermaßen umgehen zu können (Receptivity; ω_h=.83). Die Beliefs About Peer-Feedback Scale (Huisman et al., 2020) misst, als wie wertvoll Peer-Feedback als Instruktionsmethode eingeschätzt wird (α=.80), wie hoch die Zuversicht in die Qualität des eigenen Peer-Feedbacks ist (α=.78), wie hoch die Zuversicht ist, von den Peers hochwertiges Feedback zu erhalten (α=.87), und wie wertvoll die Peer-Feedback-Fähigkeiten eingeschätzt werden (α=.87). Zuletzt wurde der Feedback Perceptions Questionnaire (Strijbos et al., 2021) eingesetzt, um Fairness, Nützlichkeit und Akzeptanz des erhaltenen Feedbacks zu messen und zu einem globalen Mittelwert der wahrgenommenen Angemessenheit zu aggregieren (ω_h =.77).

Studierende haben erwartungsgemäß das Feedback, das sie von Peers auf ihre eigene Aufgabenbearbeitung erhalten haben, als umso angemessener eingeschätzt, je empfänglicher für Feedback sie sich selbst einschätzten (β=.17, *p*=.018, einseitig) und je zuversichtlicher sie waren, dass Peer-Feedback grundsätzlich hochwertig ist (β=.12, *p*=.0496, einseitig; Mehrebenen-Regression). Andere Merkmale hingen nicht signifikant mit der wahrgenommenen Angemessenheit des erhaltenen Feedbacks zusammen (alle *p*>.34). Dagegen hing die wahrgenommene Qualität des Feedbacks nicht mit den gemessenen Merkmalen derjenigen zusammen, die es geschrieben haben (alle *p*>.229; Mehrebenen-Regression). Dieses Befundmuster deutet darauf hin, dass zumindest im vorliegenden Kontext (Lehramtsstudierende höheren Semesters; Lehr-Lern-Setting erfordert

intensives Engagement bei der Feedbackgabe) die Eigenschaften der Feedback-Empfänger:innen wichtiger für die Verarbeitung des empfangenen Feedbacks als die der Feedback-Geber:innen sind: Die Offenheit, auch negatives Feedback konstruktiv zu verarbeiten, und der Glaube, dass Peer-Feedback grundsätzlich von hoher Qualität ist, helfen dabei, Feedback als angemessen wahrzunehmen und in der Folge mutmaßlich intensiver zu verarbeiten. Diese Überzeugungen könnten als Ansatzpunkt für Fördermaßnahmen dienen. Bis zur Konferenz wird auch die tatsächliche, objektive Feedbackqualität kodiert und in die Analysen einbezogen.

Werden nicht-kognitive Aspekte des Lernens von Feedbackprovidern und -personalisierung beeinflusst?

Theresa Lena Ruwe[1], Elisabeth Mayweg[2]
[1]Humboldt-Universität zu Berlin; [2]Humboldt-Universität zu Berlin

Feedback kann sich positiv auf das Lernen (z.B. die Fähigkeit Studierender zum argumentativen Schreiben (Latifi et al., 2019)) auswirken (Wisniewski et al., 2020). Dennoch zeigt etwa ein Drittel der Feedbackinterventionen negative Lerneffekte (Kluger & DeNisi, 1996), wobei die Gründe noch nicht hinreichend beforscht sind. Da Feedbackprozesse soziale Interaktionen sind (Panadero & Lipnevich, 2022), sollten soziale Aspekte, wie der Feedbackprovider (z.B. Menschen vs. Computer; Strijbos, Narciss, et al., 2010) oder die Sprache (z.B. Personalisierung; Moreno & Mayer, 2004) in Zusammenhang mit der Effektivität von Feedback gebracht werden (Ramani et al., 2019). Solche sozialen Cues rufen Social Presence (Short et al., 1976) hervor, welche durch gestiegenes Interesse und Auseinandersetzung lernfördernd wirkt (Seitchik et al., 2017). Auch nicht-kognitive Aspekte wie Selbstwirksamkeit (Bandura, 1986), Motivation (Deci & Ryan, 2012) und Emotionen (Camacho-Morles et al., 2021) spielen eine bedeutende Rolle für Lerneffekte (Schneider et al., 2021). Beide Forschungsstränge deuten also darauf hin, soziale und nicht-kognitive Faktoren vermehrt zu untersuchen, da sie die Lernerfolge beeinflussen. Darauf basierend untersucht diese Studie wie sich verschiedene Feedbackprovider und Feedbackpersonalisierung auf Selbstwirksamkeit, Motivation und Emotionen von Studierenden und damit auf ihre Überarbeitungsleistung auswirken. Wir erwarten, dass menschliche Feedbackprovider positiver wahrgenommen werden als eine KI sowie dass personalisierte im Vergleich zu neutraler Sprache nicht nur generell, sondern auch auf Motivation, Selbstwirksamkeit und Emotionen positiv wirkt. Außerdem untersuchen wir explorativ die wechselwirkenden Einflüsse von Feedbackprovider und Sprache auf die abhängigen Variablen. In einem 3 (Feedbackprovider: Peer, Lehrkraft, KI) x 2 ((nicht-)personalisiert) between-subject Design haben N = 98 Lehramtsstudierende einen argumentativen Text zum Thema Klassenführung geschrieben, Feedback bekommen und ihren Text entsprechend überarbeitet. Die online Studie griff auf valide Skalen zur Operationalisierung der abhängigen Variablen zurück: Verschiedene Facetten von Motivation wurden mit der Situational-Motivation-Scale (Guay et al., 2000) erfasst; Selbstwirksamkeit zum argumentativen Schreiben wurde basierend auf der Self-Efficacy-Writing-Scale (Bruning, 2013; Putra et al., 2020) angepasst; lernrelevante Emotionen (z.B. Freude) wurden mit dem AEQ-S (Bieleke et al., 2021) erfasst; die Erfassung der Wahrnehmung des Feedbacks (z.B. Fairness) basierte auf Strijbos, Pat-El und Narciss (2010; Feedback-Perceptions-Questionnaire), die der Feedbackprovider (z.B. Kompetenz) auf Hendriks et al. (2015; Muenster-Epistemic-Trust-

Inventory). Die Ergebnisse der Regressionsanalysen zeigen, dass eine KI als Feedbackprovider als vertrauenswürdiger als Menschen wahrgenommen wird ($F(2, 95) = 3.43$, $p < .05$, $R2 = 0.07$, $\eta2p = 0.07$) und ihr Feedback positivere Emotionen hervorruft ($F(2, 94) = 2.66$, $p < .1$, $R2= 0.05$, $\eta2p = 0.05$). Außerdem beeinflusst neutrale Sprache die Motivation ($F(3, 93) = 2.71$, $p < .05$, $R2= 0.01$, $\eta2p = 0.05$) und Selbstwirksamkeit hinsichtlich der Ideenbildung negativ ($F(7, 89) = 2.01$, $p < .1$, $R2= 0.14$, $\eta2p = 0.02$). Weitere Berechnungen zeigen, dass neutrales Feedback von Lehrenden am fairsten wahrgenommen wird ($F(7, 89) = 1.91$, $p < .01$, $R2= 0.13$, $\eta2p = 0.09$). Zudem steigt die Selbstwirksamkeit zur Ideenbildung durch neutrales Feedback einer KI ($F(7, 89) = 2.01$, $p < .1$, $R2= 0.14$ 2.261; $\eta2p = 0.04$). Dies bestätigt, dass soziale Cues in online Feedbackumgebungen eine zentrale Rolle einnehmen. Besonders hervorzuheben ist, dass eine KI gegenüber Menschen zu positiveren Effekten führt (in Bezug auf Kompetenz, Emotionen und Selbstwirksamkeit). Dennoch sollte man nicht vergessen, dass potenzielle Moderatoren wie Lehrkräfte-Einstellungen (Fernández-Batanero et al., 2021) eine Rolle spielen. Außerdem signalisieren die Ergebnisse, übereinstimmend mit der Literatur (Moreno & Mayer, 2004), Feedback personalisiert zu formulieren, um negative Effekte auf Motivation zu vermeiden. Hinsichtlich Selbstwirksamkeit sind die Ergebnisse nicht eindeutig (Ginns et al., 2013; Wang et al., 2008). Schlussfolgernd spielt Social Presence, hervorgerufen durch Feedbackprovider und -personalisierung, eine wichtige Rolle. Die Ergebnisse werden in Bezug auf ihre Implikationen für die Bildungspraxis und weitere Forschung (in Bezug auf nicht-kognitive und soziale Aspekte des Lernens) diskutiert.

Begünstigt individualisiertes Feedback lernförderliches Verhalten beim Lesen multipler Dokumente? Ein einsemestriges Moodle-Experiment

Carolin Hahnel[1,2], Daniel Bengs[1], Daniel Biedermann[1], George-Petru Ciordas-Hertel[1], Beate Eichmann[1,2], Frank Goldhammer[1,2], Ilka Wolter[3,4], Cordula Artelt[3,4]

[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; [2]Zentrum für internationale Bildungsvergleichsstudien (ZIB); [3]Leibniz-Institut für Bildungsverläufe (LIfBi); [4]Otto-Friedrich-Universität Bamberg

Das Lesen und Verstehen multipler Dokumente (MDC, *multiple document comprehension*) spielt im Studium eine besondere Rolle (z. B. Schoor et al., 2022). Dabei gestellte Anforderungen können vor allem durch zwei zentrale Prozesse bewältigt werden (Philipp, 2020): durch Informationsintegration werden Informationen dokumentenübergreifend organisiert und verknüpft (Barzilai et al., 2018) und durch Sourcing werden Metainformationen über Dokumente bewertet und berücksichtigt (Bråten et al., 2018). Die bisherige Forschungsliteratur zeigt allerdings, dass Verhaltensweisen dieser MDC-spezifischen Prozesse oft nicht spontan gezeigt werden (z. B. Wiley et al., 2009).

Eine Möglichkeit entsprechendes Verhalten von Studierenden zu fördern besteht darin, ihnen im Rahmen eines formativen Assessments konkrete Rückmeldungen zu ihrem Vorgehen beim Lesen multipler Dokumente zu geben (*learning-oriented feedback* in Johnson et al., 1993; siehe auch Schoor et al., 2022). Mit unserem Beitrag untersuchen wir, ob ein solches prozessbezogenes Feedback gewünschte MDC-spezifische Verhaltensweisen von Studierenden fördern und ihr Lernen unterstützen kann.

Zur Beantwortung unserer Forschungsfrage haben wir die Ergebnis- und Prozessdaten von 452 Lehramtsstudierenden (75.7% weiblich; M_{Alter} = 22.1 Jahre) von zwei deutschen Universitäten ausgewertet. Die Studierenden nahmen an einem viermonatigen Moodle-Kurs zum Thema „Einsatz

digitaler Medien im Unterricht“ teil. Vor und nach dem Kurs absolvierten sie einen Wissenstest. Der Kurs selbst war in fünf zweiwöchige und asynchron bearbeitbare Lerneinheiten eingeteilt, in denen vier Lerneinheiten mehrere Textdokumente enthielten. Für jede dieser Lerneinheiten wurden anhand von Logdaten verschiedene Prozessindikatoren abgeleitet und zu Gesamtscores der Informationsintegration und des Sourcing modelliert (Einheit in Prozent). Weiterhin wurden für jeden Text die Lesezeit der Studierenden erfasst. Die Prozessindikatoren stellten die Grundlage für die Erstellung eines prozessbezogenen Dashboard-Feedbacks dar.

Die Studierenden wurden zu Kursbeginn in eine von zwei Feedback-Gruppen eingeteilt. Die Kontrollgruppe (n=213) erhielt in jeder Lerneinheit ein allgemeines Feedback über das Abschneiden einer Vergleichsgruppe (z. B. „27.91% der Studierenden empfanden das akademische Lesen als eher schwierig“). Die Treatment-Gruppe erhielt ein individualisiertes Feedback zu ihrem Verhalten beim Bearbeiten der Dokumente einer Lerneinheit (n=239). Das individualisierte Feedback umfasste Rückmeldungen zu den MDC-spezifischen Prozessen in den Kategorien Beobachtung („Sie haben“), Stärken („Behalten Sie folgende Verhaltensweisen bei“) und Entwicklungsmöglichkeiten („Zusätzliche Tipps für die Bearbeitung weiterer Lerneinheiten“).

Zur Auswertung wurde der Verlauf der beobachteten Gesamtscores zur Informationsintegration und zum Sourcing über die Lerneinheiten hinweg anhand von Wachstumkurvenmodellen beschrieben. Die ermittelten Ausgangsniveaus (Interzepte) und Veränderungen (Steigungen) wurden zwischen der Treatment- und Kontrollgruppe verglichen.

Die Ergebnisse zeigten, dass die Kontrollgruppe (b=21.9, p<.001) ein durchschnittlich höheres Ausgangsniveau in Verhaltensweisen der Informationsintegration als die Treatment-Gruppe (b=16.7, p<.001) aufwies. Allerdings nahmen diese Verhaltensweisen im Semesterverlauf ab (b=-2.2, p<.01), während das Informationsintegrationsverhalten der Treatment-Gruppe zunahm (b=2.8, p<.01). Beim Sourcing waren die Ausgangsniveaus beider Gruppen sehr gering (Kontrollgruppe: b=2.1, p<.05; Treatment-Gruppe: b=3.0, p<.001). Im Gegensatz zur Kontrollgruppe (b=-0.0, n.s.) nahm das Sourcing-Verhalten in der Treatment-Gruppe allerdings über den Kursverlauf hinweg zu (b=2.5, p<.001).

Die Ergebnisse entsprechen bisherigen Befunden darüber, dass MDC-spezifische Verhaltensweise wenig spontan gezeigt werden (z. B. Wiley et al., 2009). Durch eine vergleichsweise geringfügige Intervention, wie dem individualisierten Dashboard-Feedback, konnten die gewünschten Verhaltensweise jedoch gefördert werden. Obwohl dieser Befund positiv zu werten ist, stellt sich dennoch die Frage, warum MDC-spezifisches Verhalten nicht noch stärker gezeigt wurde. Eine mögliche Erklärung könnte darin liegen, dass Studierende eventuell nicht den Eindruck gewinnen, dass MDC-spezifisches Verhalten entscheidend für ihren Lernerfolg wäre (z. B. da Texte von Dozierenden vorausgewählt werden und nicht sonderlich kritisch reflektiert werden müssen).

In weiteren Analysen werden wir noch evaluieren, ob die in der Treatment-Gruppe beobachtete Zunahme MDC-spezifischen Verhaltens auch mit einem größeren Lernerfolg einhergeht. Dazu werden das Ausgangsniveau und die Veränderung der MDC-spezifischen Verhaltensweisen

gemeinsam mit den Leistungen in den Prä- und Post-Kurs-Wissenstests betrachtet. Zum Zeitpunkt der Konferenz werden diese Analysen abgeschlossen sein und im Vortrag präsentiert werden.

G-09: Motivation und Interesse im Studium

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S05 T00 B83
Chair der Sitzung: Lena Roemer

Sollten wir uns für Berufsinteresse-States interessieren? Investigatives State Interesse in täglichen Situationen sagt Studienengagement vorher

Lena Roemer[1], Gundula Stoll[2], James Rounds[3], Matthias Ziegler[4]
[1]GESIS - Leibniz Institut für Sozialwissenschaften; [2]Universität Tübingen, Hector-Institut für Empirische Bildungsforschung; [3]University of Illinois at Urbana-Champaign, Urbana, IL, USA; [4]Humboldt-Universität zu Berlin

Berufsinteressen werden traditionell als stabile, interindividuelle Unterschiede angesehen, die Studien- oder Berufsergebnisse vorhersagen können (Holland, 1997; Rounds & Su, 2014). Beispielsweise suchen Menschen Studiengänge oder Berufe aus, die ihren Interessen entsprechen und sind zufriedener und leistungsstärker, wenn das Umfeld zu den eigenen Interessen passt (Nye et al., 2012, 2017). Ein kürzlich vorgestelltes theoretisches Rahmenmodell integriert eine momentane Perspektive in die Konzeptualisierung von Interessen (Su et al., 2019). Dieses Modell postuliert, dass sich Interesse-Traits als intraindividuell variierende States manifestieren; und dass Interesse-States über die Zeit zur Vorhersage psychologisch relevanter Outcomes und Entwicklung von Interesse-Traits kumulieren. Eine State-Perspektive auf Berufsinteressen ist also nötig, um zu untersuchen, wie sich Berufsinteressen und ihre prädiktive Kraft im täglichen Leben entfalten.

Gleichzeitig wurden Berufsinteressen auf einer State-Ebene bisher nur unzureichend untersucht. Erste Studien zeigen, dass sich Berufsinteressen als momentan variierende States manifestieren (Ziegler et al., 2018) und dass Berufsinteresse-States systematisch mit spezifischen Situationsmerkmalen zusammenhängen (Roemer et al., 2021). Inwiefern Berufsinteresse-States jedoch relevant für psychologische Outcomes sind, bleibt eine offene Frage.

Ziel dieser präregistrierten Studie ist es, die prädiktive Kraft von Berufsinteresse-States zu untersuchen. Basierend auf der Annahme, dass Menschen höhere Leistungen zeigen, wenn sie in Bereichen arbeiten, die ihren Interessen entsprechen, untersuchen wir den Zusammenhang zwischen Berufsinteresse-States und Studienengagement. Wenn sich die psychologische Relevanz von Berufsinteressen tatsächlich durch momentane States entfaltet, dann sollten Unterschiede, wie interessant Individuen ihren Studienbereich in verschiedenen Situationen finden, Unterschiede im Studienengagement vorhersagen können.

Im Rahmen dieser Experience-Sampling Studie haben N = 217 Psychologiestudierende insgesamt N_{obs} = 5631 Mal ihre Berufsinteresse-States (Rounds et al., 2016) und aktuelle Situationscharakteristika (Rauthmann et al., 2014) angegeben. Nach Abschluss der Experience-Sampling Phase haben die Studierenden ihr generelles Studienengagement (Schaufeli et al., 2006) angegeben. Die Ergebnisse zeigen, dass Studierende mit höheren investigativen Berufsinteresse-States stärkeres Studienengagement berichteten, auch nach Kontrolle für den stabilen Interesse-Trait. Dies zeigt die prädiktive Validität von Berufsinteresse-States und impliziert, dass die psychologischen Outcomes von Berufsinteressen durch eine momentane Perspektive nuancierter vorhergesagt werden können. Zudem zeigten sich schwache Effekte von Interessensreaktivität: Personen, deren momentanes investigatives Interesse eng an die Wahrnehmung von studiensbezogenen Situationen gekoppelt war, waren im Schnitt engagierter in ihrem Studium. Es

scheint also nicht nur relevant zu sein, *dass* man momentan interessiert ist, sondern auch wichtig sein, in welcher *Situation* man es ist.

Insgesamt erlaubt eine State-Perspektive also einen genaueren Blick auf die psychologische Relevanz von Berufsinteressen und führt so zu theoretischen und praktischen Implikationen.

Stark anfangen? Nicht nachlassen! Motivationsverläufe im Rahmen non-formaler Online-Kurse

Maria Klose[1], Philipp Handschuh[1], Diana Steger[2]
[1]Leibniz-Institut für Bildungsverläufe, Deutschland; [2]Universität Kassel, Deutschland

Das digitale Lernen im außerinstitutionellen Kontext wird für viele Menschen immer selbstverständlicher (Autorengruppe Bildungsberichterstattung, 2020). Bisherige Forschung zu Motivation und Persistenz von Teilnehmenden an online Lernangeboten bieten vorrangig eine querschnittliche Perspektive, da die Motivation der Teilnehmenden einmalig entweder zu Beginn oder am Ende der Studie erfasst wurde (z.B. de Barba et al., 2016). Im Fokus unserer Studie steht die längsschnittliche Betrachtung des Motivationsverlaufs bei der Nutzung einer non-formalen, digitalen Lernplattform. Dabei wird Motivation als ein dynamischer Zustand betrachtet, der variabel über die Zeit ist und sich situationsbedingt durch die Inhalte und den Kontext des non-formalen Lernens verändern kann (Guay et al., 2000). Neben dem Einfluss von situativem Interesse (Linnenbrink-Garcia et al., 2010) auf situative Motivation, betrachten wir die Selbstregulation der Lernenden, da diese Fähigkeit wesentlich für das eigenständige Lernen in Online-Kursen ist (Kizilcec et al., 2017). Darüber hinaus untersuchen wir den Zusammenhang zwischen der Lernzielorientierung und den Motivationsverläufen der Lernenden, da Lernziele die intrinsische Motivation fördern (Elliot & Church, 1997).

Ziel der Studie ist es herauszufinden, ob sich Individuen hinsichtlich ihres anfänglichen Motivationsniveaus und ihrer Motivationsverläufe über den Zeitraum der Kursbearbeitung hinweg unterscheiden (Gleichbleiben bzw. Zu- oder Abnahme der intrinsischen Motivation), um anschließend folgende Hypothesen empirisch zu testen: die intrinsische situative Motivation der Lernenden hängt positiv zusammen mit a) der Selbstregulationsfähigkeit, b) dem situativen Interesse im Laufe der Kursbearbeitung, und c) der Lernzielorientierung.

Die Datenerhebung fand von Oktober 2021 bis Juni 2022 im Rahmen des digitalen Kursangebotes der OPEN vhb (Virtuelle Hochschule Bayern, 2022) statt und umfasst bis zu drei Messzeitpunkte in 51 Online-Kursen, die ohne vorgegebene Lernzeiten, kostenfrei nutzbar und frei zugänglich waren. Die Messzeitpunkte wurden jeweils zu Beginn, nach 50% der Lektionen, sowie am Ende der Online-Kurse platziert. In allen drei Messzeitpunkten wurde *situative Motivation* (*SIMS*; Guay et al., 2000) und *situatives Interesse* (*SIS*; Linnenbrink-Garcia et al., 2010) erfasst. Zu Messzeitpunkt 1 wurde außerdem *Selbstregulation* (Subskala „Strategien des Managements interner Ressourcen“; Klingsieck, 2018) und zu Messzeitpunkt 2 *Lernzielorientierung* (Subskala von *SELLMO*; Spinath et al., 2012) erhoben. Zur Beantwortung der Frage, ob sich die Individuen in Bezug auf ihre Motivationsverläufe im Laufe der Kursbearbeitung unterscheiden, berechnen wir Wachstumskurvenmodelle. Die abhängige Variable ist *intrinsische situative Motivation*. Als Einflussvariablen betrachten wir Selbstregulation, situatives Interesse und Lernzielorientierung. Die

Stichprobe besteht aus OPEN vhb-Nutzenden, die sich freiwillig dazu entscheiden konnten an der längsschnittlichen Befragung teilzunehmen. Für die vorliegenden Analysen nutzen wir die Daten von 240 Personen, die an allen drei Messzeitpunkten teilgenommen haben (Alter in Jahren: *M* = 40.95, *SD* = 14.81, Range [18; 77]; davon 70% weiblich).

Vorläufige Ergebnisse zeigen, dass sich die Lernenden vor allem in Bezug auf ihre Anfangsmotivation unterscheiden und die intrinsische Motivation zeitlich relativ stabil ist. Die Anfangsmotivation wird positiv von situativen Interesse und Lernzielorientierung beeinflusst, nicht aber von der Selbstregulationsfähigkeit.

Ziel der Studie ist es bisherige Kenntnisse der Motivationsforschung im neuen Kontext der non-formalen digitalen Lerngelegenheiten anzuwenden. Nicht zuletzt wegen der COVID-19 Pandemie gewinnt das eigenständige Online-Lernen immer mehr an Bedeutung. Dabei scheint die situative Motivation im Rahmen non-formaler Online-Kurse zeitlich stabil zu sein und vor allem die Anfangsmotivation im Zusammenhang mit dem situativen Interesse und der Lernzielorientierung entscheidend für den weiteren Verlauf der Motivation während der Kursbearbeitung. Auf dieser Grundlage können Ansatzpunkte für individuelle Lerndesigns gefunden werden, welche die Lernenden adaptiv bei ihrer Zielerreichung unterstützen.

Zusammenhang von intrinsischer Studienmotivation und Noten über den Studienverlauf hinweg: Ein Random-Intercept Cross-Lagged Panel Modell

Laura Aglaia Sophia Messerer[1,2], Ronny Scherer[2], Stefan Janke[1]
[1]Universität Mannheim, Germany; [2]Centre for Educational Measurement, Oslo University, Norway

Studienabbruch ist ein verbreitetes Problem: Etwa 30% der Studierenden schließen ihr ursprünglich angestrebtes Studienfach nicht ab (Heublein et al., 2017; OECD, 2018). Mangelnde intrinsische Motivation und Leistungen (schlechte Noten) sind zwei Gründe hierfür—erstere in der Anfangsphase des Studiums und letztere eher in der Studienendphase. Allerdings mangelt es an Längsschnittuntersuchungen zum Zusammenspiel dieser Variablen im Studienverlauf. Gemäß der Selbstbestimmungstheorie (Deci & Ryan, 1985) entsteht bei Studierenden, die gute Leistung erbringen, das Gefühl von Kompetenz, was die intrinsische Motivation steigern kann. Erleben Studierende jedoch wiederholt Misserfolge, kann dies eher demotivieren. Die Selbstbestimmungstheorie erklärt auch, wie sich die intrinsische Motivation auf die Leistung im Studium auswirken kann: Wenn Studierende intrinsisch motiviert sind, wirkt sich das wahrscheinlich positiv auf Spaß, Anstrengungsbereitschaft und Zeitinvestment aus, was infolge zu besseren Prüfungsleistungen führt.

Um bestehende Forschungslücken bezüglich des Zusammenspiels von intrinsischer Studienmotivation und Leistung zu schließen, untersuchen wir, in welchem Ausmaß die beiden Variablen über das Studium hinweg in Beziehung stehen und ob es Studienphasen gibt, in denen die weitere Entwicklung einer der beiden Variablen maßgeblich in Beziehung zu der anderen Variablen steht.

Insgesamt wurden 1349 Bachelorstudierende einer deutschen Universität (62% weiblich, 1 nicht-binär; Durchschnittsalter zu T1 = 19,7 Jahre) über die gesamte Regelstudienzeit zu Beginn jeden Semesters befragt (6 Messzeitpunkte). Die intrinsische Studienmotivation wurde mit einer übersetzten

und an den Studienkontext angepassten Version des „Intrinsic Motivation Inventory“ (Deci & Ryan, 2013) erfasst. Die Studienleistung wurde durch die Studiennoten (Interpretationshinweis: niedrigerer Wert = bessere Leistung) ermittelt, welche von der Studienadministration übermittelt wurden. Um auszuschließen, dass Zusammenhänge zwischen intrinsischer Studienmotivation und Studienleistung Selektionseffekte reflektieren (d.h. bessere Leistung bedingt die Zulassung ins gewünschte Studienfach) wurde in dem Modell für Einflüsse der Hochschulzugangsberechtigungsnote kontrolliert.

Um die Stärke stabiler Zusammenhänge zwischen den Variablen taxierbar zu machen und gleichzeitig kritische Studienphasen zu identifizieren, in welchen diese in Wechselwirkung treten, schätzten wir ein latentes Random-Intercept Cross-Lagged Panel Modell (Hamaker et al., 2015), das die Daten gut repräsentierte, $\chi^2(730) = 1092.02$, CFI = .978, TLI = .974, RMSEA = .021, SRMR = .047. Das Model zeigte eine stabile, negative Korrelation zwischen intrinsischer Studienmotivation und dem Notendurchschnitt über die Messzeitpunkte hinweg (r = -.478, p < .001). Auf der intra-individuellen (within) Ebene zeigten sich Übertragungseffekte dahingehend, dass sowohl Leistung (T2-T5) als auch intrinsische Studienmotivation (T3-T6) selbstverstärkende Entwicklungsmuster zeigen (höhere Ausgangswerte als Grundlage für folgende positive Abweichung und umgekehrt). Außerdem hing eine vorübergehend erhöhte oder gedämpfte Studienleistung (Noten) einer Person vom vorübergehenden Motivationsniveau dieser Person zum vorherigen Messzeitpunkt ab. Dies wird durch die Kreuzeffekte zwischen intrinsischer Studienmotivation (T4) und Studiennoten (T5) indiziert, β = -.167, p = .029, sowie zwischen intrinsischer Studienmotivation (T5) und Studiennoten (T6), β = -.148, p = .030. Wenn eine Person eine höhere (niedrigere) intrinsische Studienmotivation berichtete (T4 und T5), waren die Studiennoten zum nächsten Messzeitpunkt (T5 und T6) besser (schlechter).

Insgesamt legen die Effekte nahe, dass es einerseits Zusammenhänge zwischen intrinsischer Studienmotivation und Studiennoten auf der inter-individuellen (between) Ebene gibt; es gibt also zeitliche stabile Unterschiede zwischen den Personen und auch einen zeitlich stabilen Anteil des Zusammenhangs zwischen intrinsischer Studienmotivation und Studiennoten. Andererseits gibt es gegen Ende des Bachelors Übertragungseffekte – wer motivierter ist hat im Durchschnitt bessere Noten und wer weniger motiviert ist schlechtere Noten. Auch wenn Studierende in der Endphase des Studiums vor allem aufgrund mangelnder Leistungen ihr Studium abzubrechen, spielt intrinsische Studienmotivation auch zu diesem Zeitpunkt eine Rolle, da sie mit der Studienleistung zusammenhängt. Intrinsische Studienmotivation spielt demnach nicht nur für selbstgewählten Studienabbruch eine Rolle, sondern auch für Studienabbruch aufgrund mangelnder Leistungen. Für die Bildungspraxis bedeutet das, dass Motivationsförderung wichtig ist, um positive Verstärkungseffekte anzustoßen bzw. dem Teufelskreis aus immer geringerer Motivation und damit geringeren Leistungen entgegenzuwirken.

Studienmotivation und Studienerfolg –Zusammenhänge zwischen Erwartung, Wert und motivationalen Kosten mit Studienleistungen und Studienabbruchintentionen

Theresa Schnettler[1,2], Anne Scheunemann[3], Lisa Bäulke[4], Daniel O. Thies[3], Lena S. Kegel[2], Julia Bobe[5], Markus Dresel[6], Stefan Fries[7], Detlev Leutner[8], Joachim Wirth[3], Carola Grunschel[2]

[1]Universität Mannheim; [2]Westfälische Wilhelms-Universität Münster; [3]Ruhr-Universität Bochum; [4]Universität Tübingen; [5]Universität Paderborn; [6]Universität Augsburg; [7]Universität Bielefeld; [8]Universität Duisburg-Essen

Aktuelle Studien unterstreichen die hohe Relevanz der Studienmotivation für den Studienerfolg im Sinne guter Studienleistungen und niedriger Studienabbruchintentionen (Schneider & Preckel, 2017; Wigfield et al., 2021). Studienmotivation ist ein multidimensionales Konstrukt, welches sich aus Perspektive der etablierten Erwartungs-Wert Theorie (Eccles & Wigfield, 2020) aus den Erfolgserwartungen der Studierenden und dem Wert des Studiums ergibt. Der Wert des Studiums setzt sich aus positiven (intrinsischer Wert, Wichtigkeit, Nützlichkeit) und negativen (Anstrengung, emotionale Kosten, Opportunitätskosten) Komponenten zusammen. Bisherige Studien zeigen, dass die Erfolgserwartung stark mit guten Studienleistungen zusammenhängt (Wu et al., 2020), während positive und insbesondere negative Wertkomponenten mit Studienabbruchintentionen assoziiert sind (Schnettler et al., 2020). Mehrheitlich wurden bisher isolierte Effekte motivationaler Variablen auf Studienerfolgsindikatoren betrachtet. Jedoch bedingen sich gemäß Erwartungs-Wert Theorie Erfolgserwartungen und Wertkomponenten in ihrem Einfluss auf den Studienerfolg gegenseitig. Dies erfordert die Berücksichtigung von Interaktionseffekten zwischen Erfolgserwartungen und Wertkomponenten in den Analysen (Trautwein et al., 2012). Zudem kann die Erklärungskraft und Genauigkeit der Vorhersage von Studienerfolgsindikatoren durch eine weitere Differenzierung der Wertkomponenten gesteigert werden (Gaspard et al., 2015). Dazu fehlt bislang ein differenziertes Messinstrument im Hochschulkontext. Insgesamt betrachtet könnten die ermittelten Befunde das theoretische Verständnis zum Zusammenwirken motivationaler Variablen in der Vorhersage des Studienerfolgs erweitern und Hinweise zur Gestaltung von Interventionen zur Steigerung des Studienerfolgs bieten.

Der vorliegende Beitrag verfolgt drei Ziele. Das Hauptziel ist (1) die Untersuchung von Zusammenhängen zwischen Erfolgserwartungen, Wert- und Kostenkomponenten auf Studienleistungen und Studienabbruchintentionen unter (2) Hinzunahme von Interaktionseffekten der motivationalen Variablen. Darüber hinaus wird (3) eine neue Skala zur differenzierten Erfassung der Studienmotivation im Hochschulkontext vorgestellt und hinsichtlich ihrer Faktorstruktur und Güte geprüft.

Insgesamt nahmen *N* = 1435 Studierende verschiedener Studienfächer an einer Studie mit dreizehn Messzeitpunkten teil, die sich über sechs Semester (Herbstsemester 2018/2019 bis Sommersemester 2021) erstreckte. Im aktuellen Beitrag berichten wir relevante Daten vom Beginn (T1) und Ende (T3) des Wintersemesters 2018/2019. Zu T1 wurde die Studienmotivation mit drei Items zu Erfolgserwartungen (Fleischer et al., 2019) und acht Subskalen zu je drei Items zur Erfassung von Wertkomponenten in Anlehnung an Gaspard et al. (2015) erhoben. Am Ende des Semesters (T3) wurden Prüfungsnoten und Studienabbruchintentionen (Bäulke et al., 2021) erfasst.

Zur Überprüfung der Faktorstruktur und Reliabilität der neuen Skala verwendeten wir aufgrund der hierarchischen Struktur und der zahlreichen Komponenten des Werts (Part et al., 2020) ein bifaktorielles exploratives Strukturgleichungsmodell mit zufriedenstellenden Ergebnissen. Zur Untersuchung der Effekte der motivationalen Variablen auf die Studienerfolgsindikatoren wurden jeweils einzelne latente Strukturgleichungsmodelle berechnet. Zur Untersuchung der Wechselwirkungen wurden die bestehenden Modelle jeweils um latente Interaktionen zwischen Erfolgserwartungen und Wertkomponenten erweitert (Meyer et al., 2019; Klein & Moosbrugger, 2000).

Hinsichtlich der Zusammenhänge der einzelnen Prädiktoren mit Studienerfolgsindikatoren zeigte sich, dass gute Studienleistungen zu T3 mit höherer Wichtigkeit guter Leistungen und geringeren emotionalen Kosten und Opportunitätskosten zu T1 einhergingen. Hohe Studienabbruchintentionen zu T3 zeigten sich bei geringem allgemeinen Wert des Studiums und hohen emotionalen Kosten sowie hohen Opportunitätskosten zu T1. Die Untersuchung latenter Interaktionen ergab darüber hinaus negative Interaktionen der Kosten sowohl mit der Erfolgserwartung als auch mit positiven Wertkomponenten. Dabei verringerten motivationale Kosten die positiven Auswirkungen der Erfolgserwartung und der positiven Wertkomponenten auf die Studienerfolgsindikatoren.

Zusammengefasst zeigte sich also, dass Studienleistungen stark mit der spezifischen Wichtigkeit guter Leistungen assoziiert waren, während sich eine globalere Einschätzung des Werts des Studiums als wichtig für Abbruchintentionen erwies. Zudem beeinflusste das Erleben von motivationalen Kosten beide Studienerfolgsindikatoren negativ (Jiang et al., 2018). Zugleich verminderten motivationale Kosten den positiven Einfluss von Erfolgserwartungen und positiven Wertkomponenten auf die Studienerfolgsindikatoren. Interventionen sollten daher neben Aspekten zur Steigerung der Motivation auch Ansätze zur Reduktion motivationaler Kosten beinhalten (Rosenzweig et al., 2022). Zukünftige Studien könnten zudem zwischen Studienfächern differenzieren, um kontextuelle Einflüsse auf die Studienmotivation zu beleuchten.

G-10: Expertise und Überzeugungen von Lehrkräften

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S05 R03 H20
Chair der Sitzung: Julia Schiefer

Machen die Lehrkräfte einen Unterschied? Die Relevanz individueller Charakteristika Lehrender in einem außerunterrichtlichen Lernsetting für Grundschulkinder

Julia Schiefer[1], Tim Fütterer[2], Wolfgang Wagner[2]
[1]Christian-Albrechts-Universität zu Kiel; [2]Eberhard Karls Universität Tübingen

Theoretischer Hintergrund und Fragestellung

Außerunterrichtliche Lernorte wie Schülerlabore oder Enrichment Programme bieten besondere Lernsettings, um Schüler*innen über die Lerninhalte des Regelunterrichts hinaus individuell zu fördern (z.B. Golle et al., 2017). Gleichwohl Lehrende zentral für die Lernwirksamkeit solcher Lernsettings sind (z.B. Stake & Mares, 2001), existiert in diesem Kontext bislang wenig Forschung zu der Relevanz individueller Charakteristika Lehrender. Abgeleitet aus einschlägiger Forschung zu Lehrenden im Schulkontext ist davon auszugehen, dass deren Professionswissen, Motivation und Überzeugungen mit den Lernerfolgen der Lernenden in Zusammenhang stehen (Blömeke et al., 2015; Kunter et al., 2011). Die vorliegende, präregistrierte Studie untersuchte diese Zusammenhänge im Rahmen eines außerunterrichtlichen Kurses zur Förderung des wissenschaftlichen Denkens von Dritt- und Viertklässler*innen, der seit 2015 fester Bestandteil eines Enrichmentprogramms in Baden-Württemberg ist. Der zehnwöchige Kurs fokussiert die Förderung des Wissenschaftsverständnisses—insbesondere Vorstellungen zur Natur des Wissens und des Wissenserwerbs—und der Motivation für Naturwissenschaften von Kindern (Autor, 2017, 2020). Konkret wurde untersucht, ob zwischen den individuellen Charakteristika der Lehrenden und dem Lernzuwachs der Kinder im Kurs positive Zusammenhänge bestehen. Wir erwarteten, dass (1) die epistemischen Überzeugungen der Lehrenden mit den epistemischen Überzeugungen der Kinder, (2) die Begeisterung für das Unterrichten mit dem Wissenschaftsverständnis und der Motivation der Kinder sowie (3) die Lehrüberzeugungen (d.h. Überzeugungen zu kognitiv aktivierendem Unterricht) mit dem Lernzuwachs der Kinder zusammenhängen.

Methode

Stichprobe und Design

Es wurden Daten aus zwei randomisierten, kontrollierten Feldstudien verwendet. Insgesamt nahmen 427 Kinder teil, die sich in der dritten und vierten Klasse befanden (Kohorte 1: 117 Kinder, 10 Lehrende; Kohorte 2: 310 Kinder, 30 Lehrende). Die Lehrenden (42,5 % männlich; Alter: M = 44.43, SD = 13.00) waren sowohl Lehrkräfte als auch Expert*innen in verschiedenen naturwissenschaftlichen Fächern und nahmen vor der Kursdurchführung an einer eintägigen Schulung teil.

Instrumente

Zur Erfassung des Wissenschaftsverständnisses der Kinder wurde der SIC-Test (scientific inquiry cycle, Schiefer et al., 2019) eingesetzt, epistemische Überzeugungen wurden mit den Skalen „Sicherheit“ und „Entwicklung des Wissens“ von Conley et al. (2004) und das intrinsische Interesse am Forschen mit einem Instrument in Anlehnung an Gaspard et al. (2017) erfasst. Bei den Lehrenden wurden epistemische Überzeugungen (Conley et al., 2004), Enthusiasmus für das Unterrichten

(selbstgeneriert basierend auf COACTIV-R) sowie genetisch-sokratische Lehrüberzeugungen als Maß für Überzeugungen zu kognitiv aktivierendem Unterricht erfasst (nach Baumert et al., 2008). Kovariaten waren die jeweiligen Prätestwerte, kognitive Fähigkeiten, Geschlecht, Alter und Schulnoten (Kinder) sowie Lehrerfahrung (Lehrende). Die Durchführungstreue des Kurses wurde im Selbstbericht dokumentiert und war zufriedenstellend (>90%).

Statistische Analyse

Für alle eingesetzten Messinstrumente wurden konfirmatorische Faktorenanalysen gerechnet. Die Zusammenhänge zwischen den Schüler- und Lehrervariablen wurden mittels multipler Regressionsanalysen in Mplus (AVs: standardisierte Posttestwerte der Schüler*innen) analysiert. Einseitige Signifikanztests werden berichtet.

Ergebnisse und ihre Bedeutung

Epistemische Überzeugungen der Lehrenden hingen positiv mit den epistemischen Überzeugungen zur Sicherheit des Wissens der Kinder zusammen (Entwicklung Lehrende: B=.19, p=.003; Sicherheit Lehrende: B=.16, p=.011). Zudem zeigte sich ein positiver Zusammenhang zwischen dem Enthusiasmus der Lehrenden und dem Verständnis der Kinder für den SIC (B=.13, p=.026). Ein negativer Zusammenhang zeigte sich hingegen zwischen den genetisch-sokratischen Lehrüberzeugungen und der Motivation der Kinder (B=-.18, p=.010).

Die Ergebnisse liefern Hinweise darauf, dass sich epistemische Überzeugungen von Lehrenden auf die Entwicklung von epistemischen Überzeugungen von Kindern im Rahmen eines solchen Kurses auswirken und somit eine wichtige Voraussetzung für die Förderung von Vorstellungen zu Naturwissenschaften darstellen. Auch der Enthusiasmus der Lehrenden hat einen Einfluss auf die Entwicklung naturwissenschaftlicher Konzepte der Kinder. Die Untersuchung gelingender Lehr-Lernprozesse in diesem Bereich ist insbesondere in Zeiten zunehmender Informationsfluten entscheidend, um Schüler*innen dabei zu unterstützen wissenschaftliche Evidenzen einordnen und bewerten zu können. Die Einstellungen zu kognitiver Aktivierung scheinen hingegen einen negativen Effekt (ggf. eine Überforderung) auf das Interesse der Kinder zu haben.

Expertisebedingte Unterschiede in der professionellen Wahrnehmung von frühpädagogischen Fachkräften – eine Eyetracking-Studie

Jennifer Busch[1], Hendrik Lohse-Bossenz[2]
[1]Pädagogische Hochschule Heidelberg; [2]Universität Greifswald

Theoretischer Hintergrund

Die professionelle Wahrnehmung und Reflexion pädagogischen Handelns sind zentrale Bestandteile in der Professionalisierung von pädagogischen Fachkräften (Mischo et al., 2020; Seidel & Stürmer, 2014). Die Professionelle Wahrnehmung als die Fähigkeit, lernrelevante Ereignisse im Unterricht zu erkennen und unterrichtliches Handeln daraufhin auszurichten, beinhaltet zwei wissensbasierte Prozesse (noticing und knowledge-based reasoning), die verschiedene Fähigkeiten erfordern (Seidel et al., 2010). Während die Konzepte der professionellen Wahrnehmung und Reflexion im Bereich der Lehrkräfteforschung bereits umfassend untersucht wurden, liegen für das Konzept der professionellen Wahrnehmung in der frühen Bildung noch keine vergleichbaren Befunde vor (Mischo et al., 2020). Da

von einer dem Lehrkräftekontext ähnlichen Bedeutung des Konzepts ausgegangen wird, werden Befunde aus dem Bereich der Lehrkräftebildung auf die Frühpädagogik übertragen.

Befunde aus der Expertiseforschung hinsichtlich der professionellen Wahrnehmung zeigen, dass erfahrene Lehrkräfte lernrelevante und komplexe Unterrichtssituationen sensitiver wahrnehmen (Gold et al., 2013). Anhand von Blickbewegungen lassen sich expertisebedingte Unterschiede in der visuellen Aufmerksamkeitsfokussierung dahingehend aufzeigen, dass Expert:innen ihre Aufmerksamkeit gezielter auf relevante Merkmale richten und diese häufiger fixieren (z.B. Stürmer et al., 2017). Zudem erkennen sie diese zuverlässiger und gehen fokussierter bei der Betrachtung einer Situation vor. Des Weiteren Auch rücken Lernende mit zunehmender Expertise stärker in den Aufmerksamkeitsfokus von pädagogischen Fachkräften (Buitink, 2009).

Fragestellung

Die vorliegende Studie geht der Frage nach, wie frühpädagogische Fachkräfte in Abhängigkeit ihrer beruflichen Expertise (operationalisiert über Berufserfahrung) Aspekte von Bildungsgelegenheiten (a-priori definierte Interessenbereiche: Situationsgestaltung, Handeln der Fachkraft und Handeln der Kinder) wahrnehmen (erfasst über Blickbewegungen) und diese reflektieren (erfasst über schriftliche Reflexionen).

Es wird erwartet, dass berufserfahrene Fachkräfte (> 4 Jahre) die Bildungsgelegenheit dahingehend wahrnehmen, dass sie ihren visuellen Aufmerksamkeitsfokus stärker auf das Handeln der Kinder richten. Dies soll anhand von höheren Fixationshäufigkeiten sowie einer höheren Fixationsdauer hinsichtlich der definierten Interessenbereiche aufgezeigt werden. Zudem wird erwartet, dass Noviz:innen eher die Situationsgestaltung sowie das Handeln der Fachkraft fokussieren. Ein ähnliches Muster wird in den schriftlichen Reflexionen erwartet.

Methode

N=16 frühpädagogische Fachkräfte (88.8% weiblich; Alter *M*=30.58 Jahre, *SD*=8.23 Jahre; Berufserfahrung *M*=5.64 Jahre, *SD*=5.59 Jahre), *n*=7 Expert:innen (> 4 Jahre Berufserfahrung) und *n*=9 (< 4 Jahre Berufserfahrung), nahmen an der explorativen Eyetracking-Studie teil. Sie betrachteten drei kurze Videos, die die Interaktion zwischen einer Fachkraft und einer Gruppe von Kindern zeigen, und reflektierten diese schriftlich. Während der Videobetrachtung wurden die Blickbewegungen der frühpädagogischen Fachkräfte aufgezeichnet. Die Datenauswertung erfolgte hinsichtlich gängiger Blickbewegungsmaße wie Fixationshäufigkeit und Fixationsdauer (Holmqvist & Andersson, 2017), die auf a-priori definierte Interessenbereiche (Situationsgestaltung, Handeln der Fachkraft & Handeln der Kinder) in den Videos ausgerichtet waren, sowie einer qualitativen Inhaltsanalyse der Reflexionstexte (Brandtner & Lohse-Bossenz, 2017), um eine Übereinstimmung zwischen visueller Betrachtung und schriftlicher Reflexion zu ermitteln. Die Interrater-Reliabilität für 586 kodierte Einheiten liegt bei Cohen‘s Kappa von κ=.83.

Ergebnisse

Erste Ergebnisse zeigen, dass alle in den Videos definierten Interessenbereiche von beiden Gruppen visuell betrachtet und schriftlich reflektiert wurden. Über alle Videos hinweg zeigt sich, dass

Noviz:innen im Vergleich zu Expert:innen die Situationsgestaltung signifikant häufiger (Fixationshäufigkeit in Anzahl; M_{Nov}=140.53, SD_{Nov}=5.37, M_{Exp}=119.88, SD_{Exp}=7.37, p=<.001und länger (Fixationsdauer in Sekunden; M_{Nov}=53.68, SD_{Nov}=1.02, M_{Exp}=44.64, SD_{Exp}=1.98, p=<.001) visuell fixieren. Bei den anderen Bereichen zeigen sich keine signifikanten Effekte. Das Antwortmuster deutet jedoch darauf hin, dass Noviz:innen ihre Aufmerksamkeit häufiger (M_{Nov}=42.10, SD_{Nov}=31.69, M_{Exp}=36.61, SD_{Exp}=26.71, p=0.713) und länger (M_{Nov}=14.19, SD_{Nov}=12.61, M_{Exp}=12.52, SD_{Exp}=8.64, p=.758) auf das Handeln der Fachkraft richten, während hingegen Expert:innen häufiger (M_{Nov}=83.50, SD_{Nov}=22.02, M_{Exp}=88.99, SD_{Exp}=12.94, p=.544) und länger (M_{Nov}=27.00, SD_{Nov}=6.81, M_{Exp}=30.05, SD_{Exp}=4.58, p=.304) das Handeln der Kinder fokussieren. Anhand der schriftlichen Reflexionen wird deutlich, dass das Handeln der beteiligten Personen unabhängig von der beruflichen Expertise häufiger thematisiert wurde. Weitere Analysen sollen die expertisebedingten Unterschiede näher beleuchten. Zudem soll untersucht werden, wie Fachkräfte vor dem Hintergrund ihrer beruflichen Expertise bei der Betrachtung von videografierten Situationen vorgehen.

Bezugsnormmuster von Lehrkräften inklusiver Lehr-Lernkontexte

Josephine Tanneberger, Jessica Jaeuthe, Sebastian Kempert, Nadine Spörer
Universität Potsdam, Germany

Konzeptueller Hintergrund

Im Rahmen ihrer beruflichen Tätigkeit stehen Lehrkräfte einer Reihe unterschiedlichster Anforderungen gegenüber. Die Beurteilung der Lernleistungen von Schülerinnen und Schülern zählt zu den Kernaufgaben des Lehrkräfteberufs (Kultusministerkonferenz, 2000). Um Beurteilungspraktiken von Lehrkräfte beschreiben und erklären zu können befasst sich die bildungswissenschaftliche Forschung mit sogenannten Bezugsnormen bzw. Bezugsnormorientierungen von Lehrkräften und deren Relevanz für das schulische Lernen (Köller, 2004; Krawinkel, Südkamp & Tröster, 2017; Rheinberg, 1980; Stang & Urhahne, 2016; Wilbert & Gerdes, 2009). Bezugsnormen bilden den Vergleichsmaßstab der benötigt wird, um eine Leistung hinsichtlich ihrer Qualität kategorisieren zu können (Rheinberg, 1980). Leistungen können dabei mithilfe einer individuellen, sozialen oder kriterialen Bezugsnorm eingeordnet werden (Heckhausen, 1974). Es zeigt sich, dass sich Lehrkräfte hinsichtlich ihrer favorisierten Bezugsnormen unterscheiden (Mischo & Rheinberg, 1995; Rheinberg, 1982). Zudem bestätigen empirische Untersuchungen, dass Bezugsnormen in der Praxis miteinander kombiniert werden und dabei einzelne Bezugsnormen dominieren (Holder & Kessels, 2019; Holmeier, 2013; Lintorf & Buch, 2021). Es deutet sich weiterhin an, dass die Bezugsnormwahl von Lehrkräften abhängig von Beurteilungszielen (Lintorf & Buch, 2021; Schöne, Dickhäuser, Spinath & Stiensmeier-Pelster, 2004) sowie vom situativen Kontext der Beurteilung ist (Holder & Kessels, 2019; Holmeier, 2013; Rheinberg, 2017; Wilbert & Gerdes, 2009). Auszugehen ist demnach zum einen von relativ stabilen und dominanten Strukturen (Lintorf & Buch, 2021) bei einem gleichzeitigen Auftreten einer situativen Flexibilität der Bezugsnormwahl (Holmeier, 2013; Lintorf & Buch, 2021; Wilbert & Gerdes, 2009).

Bezugsnormen von Lehrkräften werden bisher oftmals eindimensional und vergleichend hinsichtlich ihres Auftretens sowie ihrer Auswirkungen auf Schüler*innenmerkmale betrachtet. Es ist wenig erforscht, wie sich die Bezugsnormwahl von Lehrkräften unter Berücksichtigung variabler

Beurteilungskontexte darstellen und typisieren lässt. Zudem ist nicht bekannt, ob sich Lehrkräfte anhand unterschiedlicher Tendenzen in ihrer Bezugsnormwahl gruppieren lassen. Der Beitrag verfolgt das Ziel inhaltlich differenzierbare Bezugsnormmuster aufzudecken.

Fragestellungen und Hypothesen

Basierend auf der Annahme der praktischen Verflechtung von Bezugsnormen sowie des Vorhandenseins stabiler, dominanter aber situationsspezifisch flexibler Strukturen bei der Bezugsnormwahl formulieren wir folgende Frage und Hypothesen:

Lassen sich bezugsnormspezifische Lehrkräftegruppen finden, die sich in Hinblick auf dominante und flexible Strukturen bei der Bezugsnormwahl unterscheiden?

H1: Unter Berücksichtigung unterschiedlicher bezugsnormrelevanter Beurteilungskontexte lassen sich Gruppen identifizieren, die sich in der Flexibilität der Bezugsnormwahl unterscheiden.

H2: Unter Berücksichtigung unterschiedlicher Bezugsnormtendenzen lassen sich Gruppen identifizieren, die je dominante Strukturen der sozialen und individuellen Bezugsnorm aufweisen.

Methode

Für die empirische Überprüfung der Hypothesen werden Querschnittsdaten aus einer Erhebung zum inklusiven Lernen im Bundesland Brandenburg (Lenkeit, Hartmann, Ehlert, Knigge & Spörer, 2021) aus dem Schuljahr 2019/20 verwendet. Einbezogen werden die Befragungsdaten von N = 2763 Regellehrkräften (92%) sowie Sonderpädagog*innen (8%) von insgesamt 214 inklusiven Schulen. Teil nahmen Lehrende aller Jahrgangsstufen in Grund- (71 %) und weiterführenden Schulen (34 %). Die Bezugsnormwahl der Lehrkräfte wurde formativ operationalisiert. Die drei Bezugsnormen werden mit je fünf situationsspezifischen Einzelindikatoren gemessen. In die Analysen fließen die 15 Einzelindikatoren sowie je ein Summenwert für jede Bezugsnorm ein. Für die Exploration möglicher Gruppen der Bezugsnormwahl in Abhängigkeit von unterschiedlichen Beurteilungskontexten (H1) werden Latente Klassenanalysen (LCA) in MPlus durchgeführt. Zudem wird über latente Profilanalysen (LPA) geprüft, ob sich dominante Strukturen der sozialen und individuellen Bezugsnorm (H2) in spezifischen Bezugsnormprofilen aufdecken lassen.

Ergebnisse und Diskussion

Die bisherigen Analysen zeigen unterschiedliche Gruppierungsmöglichkeiten von Lehrkräften anhand ihrer situativen und präferierten Bezugsnormwahl auf. Die Ergebnisdarstellung vergleicht den Datenfit der verschiedenen Klassen- (LCA, H1) und Profillösungen (LPA, H2) unter Berücksichtigung zentraler Maße des relativen Modellfits (AIC, BIC und sample size adjusted BIC). Verglichen werden zudem die Werte der Bootstrap-Likelihood-Ratio-Differenztests sowie der Vuong-Lo-Mendell-Rubin-Tests für die verschiedenen Modelle. Die inhaltlich differenzierten Lehrkräftegruppierungen werden mit Hinblick auf Implikationen für die inklusive pädagogische Praxis insbesondere vor dem Hintergrund der vielfältigen Herausforderungen bei der Berücksichtigung individueller Ausgangslagen des Lernens und Leistens diskutieren.

Auslandsaufenthalte von Lehramtsstudierenden und ihre Beziehung zu den Überzeugungen zur kulturellen Vielfalt im Klassenzimmer

Jana Costa[2], Sebastian Franz[2], Claudia Menge[1]
[1]Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW); [2]Leibniz-Institut für Bildungsverläufe (LIfBi)

Theoretischer Hintergrund

Um Lehrkräfte auf den Unterricht in heterogenen Klassen vorzubereiten und die Entwicklung (inter-)kultureller Kompetenzen zu fördern, wird häufig eine Internationalisierung der Lehrkräfteausbildung gefordert (Kissock & Richardson, 2010). Dabei wird davon ausgegangen, dass einschlägige Auslandserfahrungen zu positiven Veränderungen in den Überzeugungen und im Umgang mit Heterogenität im Klassenzimmer führen. Die Überzeugungen zählen zu den motivationalen Orientierungen von Lehrkräften und stellen damit eine relevante Facette ihrer professionellen Kompetenz dar (Baumert und Kunter 2006).

Da Auslandsaufenthalte direkte, indirekte und physiologisch-emotionale Erfahrungen ermöglichen, können sie als Stimuli zur Veränderung von Überzeugungen von Lehrkräften betrachtet werden. In Anlehnung an Helmkes (2015) Angebots-Nutzungs-Modell, das von Kunter et al. (2011) auch auf die Entwicklung der beruflichen Kompetenz von Lehrkräften übertragen wurde, werden Auslandsaufenthalte in diesem Beitrag als Lerngelegenheiten für Lehramtsstudierende beschrieben. Neben den Lernangeboten und deren Nutzung werden dabei auch die persönlichen Voraussetzungen der Lehrkraft und der Kontext der Lernumgebung als zentrale Faktoren für die Kompetenzentwicklung der Lehrkräfte angesehen.

Fragestellung

Die vorliegende Studie schließt an Forschung zur internationalen Mobilität von Lehramtsstudierenden an. Während bisherige Studien zur Mobilität von Studierenden häufig keinen klaren Fokus auf angehende Lehrkräfte setzen (vgl. Netz et al. 2020) oder lediglich einzelne Förderprogramme evaluieren (vgl. Engel et al. 2009), wird in diesem Beitrag explizit die Zielgruppe der Lehrkräfte unter Rückgriff auf ein deutschlandweites Sample in den Blick genommen.

Neben der Untersuchung der individuellen Merkmale von Lehramtsstudierenden, die ins Ausland gehen, wird der Zusammenhang zwischen Auslandserfahrungen und den (Selbstwirksamkeits-)Überzeugungen von Lehrkräften in Bezug auf kulturelle Vielfalt im Unterricht analysiert. Anschließend wird geprüft, wie robust dieser Zusammenhang zwischen Auslandsstudium und Überzeugungen ist, indem für individuelle Merkmale kontrolliert wird.

Mit dem vorgeschlagenen Beitrag wird eine klare Konzeptualisierung der (Selbstwirksamkeits-)Überzeugungen von Lehrkräften in Bezug auf kulturelle Vielfalt im Klassenzimmer angewendet und der Zusammenhang von lehrkräftespezifischen Überzeugungen mit Auslandserfahrungen während des Studiums exploriert.

Methode

Die vorliegende Studie verwendet Daten der Startkohorte 5 des Nationalen Bildungspanels (NEPS), die seit dem Wintersemester 2010/2011 eine bundesweite Zufallsstichprobe von Studienanfängerinnen und -anfängern an deutschen Hochschulen durch das Studium und ins Berufsleben begleitet (Blossfeld et al., 2011; doi:10.5157/NEPS:SC5:16.0.0).

Die Analysestichprobe der vorliegenden Studie umfasst n=1.860 (angehende) Lehrkräfte, die an der fünften Online-Befragung vom Herbst 2016 teilnahmen, als die Überzeugungen gegenüber Schüler*innen mit Migrationshintergrund gemessen wurden. Zu diesem Zeitpunkt waren etwa 45 % in einem Lehramtsstudium eingeschrieben, 40 % befanden sich im Vorbereitungsdienst und 15 % arbeiteten bereits als Lehrkraft.

Die Einstellungen gegenüber Schüler*innen mit Migrationshintergrund und kultureller Vielfalt im Klassenzimmer wurden mit einer Kurzversion der Teacher Cultural Belief Scale (Hachfeld et al., 2011) gemessen, die zwischen multikulturellen (α = ,75) und egalitären Überzeugungen (α = ,80) unterscheidet. Die Selbstwirksamkeitserwartung in Bezug auf das Unterrichten in kulturell vielfältigen Klassen wurde mithilfe einer verkürzten Skala mit drei Items (α = ,91) aus der Studie COACTIV-R (Teacher Candidates' Acquisition of Professional Competence During Teaching; Max-Planck-Institut für Bildungsforschung, 2010) erhoben.

Es werden Strukturgleichungsmodelle genutzt, um Zusammenhänge zwischen den Variablen zu analysieren.

Ergebnisse und ihre Bedeutung

Entsprechend der kurrikularen Vorgaben ist die Wahrscheinlichkeit, während des Studiums eine Auslandserfahrung gemacht zu haben, bei Personen, die ein sprachliches (Unterrichts-)Fach studiert haben, signifikant ($p < ,05$) höher. Zudem besteht ein positiver Zusammenhang mit dem Persönlichkeitsmerkmal der Verträglichkeit. Überraschenderweise wurde kein Zusammenhang zu den soziodemografischen Merkmalen wie Geschlecht, Migrationshintergrund und akademischem Hintergrund festgestellt.

In Übereinstimmung mit den theoretischen Annahmen steht ein Auslandsaufenthalt während des Studiums in einem positiven Zusammenhang mit multikulturellen und egalitären Überzeugungen sowie mit der Selbstwirksamkeitserwartung. Die beobachteten Zusammenhänge bestehen weiterhin, wenn zusätzlich Persönlichkeitseigenschaften kontrolliert werden. Offenheit für neue Erfahrungen ist sowohl mit multikulturellen als auch mit egalitären Überzeugungen positiv assoziiert. Verträglichkeit ist mit allen Outcomes positiv assoziiert.

Im Rahmen der Diskussion werden die Implikationen der Ergebnisse für die Lehrkräftebildung herausgearbeitet.

Symposien

G-11: Von Wissen zu Performanz von Lehrkräften: Beiträge zur Erfassung von Kompetenzfacetten im integrativen Kompetenzverständnis

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* S07 S00 D07

Von Wissen zu Performanz von Lehrkräften: Beiträge zur Erfassung von Kompetenzfacetten im integrativen Kompetenzverständnis

Chair(s): **Matthias Baumgartner** (Pädagogische Hochschule St.Gallen (PHSG), Switzerland)
*Diskutant*in(nen):* **Johannes König** (Universität zu Köln)

Innerhalb der kompetenzorientierten Lehrer*innenbildungsforschung hat sich in den vergangenen Jahren ein integratives Kompetenzverständnis etabliert (Blömeke et al., 2015), in welchem die analytische (z.B. Baumert & Kunter, 2006), die hybride (z.B. van Es & Sherin, 2008) sowie die holistische (z.B. Oser, 2001) kompetenzorientierte Traditionen verbunden werden (Baumgartner, 2017). Dabei konstituiert sich das Kompetenzkonstrukt aus den drei zusammenhängenden Kompetenzfacetten der a) Aspekte professioneller Kompetenz (z.B. Professionswissen), b) der situierten Wahrnehmungs-, Interpretations- und Entscheidungsprozesse (WIE) und c) der Performanz (Baumgartner, 2022). Die angenommenen Zusammenhänge konnten in Studien vereinzelt herausgestellt werden, z.B. professionelles Wissen und WIE (Kersting et al., 2010; Meschede et al., 2017) oder WIE und Unterrichtsqualität (König & Kramer, 2016, Roth et al., 2011). Forschung bezüglich pädagogischer Inhalte wie zum Beispiel Klassenmanagement zeigten dabei insbesondere mit Blick auf die Zielvariable der Performanz inkonsistente Befunde (z.B. Junker et al., 2021; König et al., 2021). Deshalb gilt es im Rahmen des (disziplinenübergreifenden) Symposiums zunächst als relevant, einer validen Erfassung der jeweiligen Kompetenzfacetten bei (angehenden) Lehrkräften nachzugehen sowie darauf aufbauenden die modellbezogenen Zusammenhangsannahmen zu untersuchen.

Im Spezifischen wird in den beiden ersten Beiträgen ein vom Schweizerischen Nationalfonds gefördertes Forschungsprojekt zum Transformationsprozess von Wissen zu Performanz bei (angehenden) Sportlehrkräften am Beispiel der Klassenführung (WiPe-Sport) vorgestellt. Im ersten Beitrag werden die Ergebnisse zur inhaltlichen und empirischen Validierung eines klassenführungsbezogenen Wissenstests bei (angehenden) Sportlehrkräften präsentiert und diskutiert. Im zweiten Beitrag werden die Resultate der inhaltlichen und empirischen Validierung des videobasierten Tests zur Erfassung der Qualität der klassenführungsbezogenen WIE dargestellt.

Der dritte Beitrag präsentiert die Entwicklung und Validierung von Testinstrumenten zur Erfassung fachspezifischen professionellen Wissens und Könnens von Sportlehrpersonen im Rahmen des vom Schweizerischen Nationalfonds geförderten Projekts EPiC-PE 3:1. Dabei wurden neben klassischen MC-Tests zur Erfassung dispositionalen Wissens auch video- und textbasierte Instrumente eingesetzt, um über die Darstellung von Unterrichtssituationen Wahrnehmungs-, Interpretations- und Entscheidungsprozesse und damit fachliche sowie fachdidaktische Wissensanwendungen bei (angehenden) Sportlehrkräften erheben zu können. Im vierten Beitrag wird der Fokus auf die differente Qualität der WIE bei Masterstudierenden des Lehramtes verschiedener Unterrichtsfächer gerichtet. Es wird die Fragestellung bearbeitet, ob sich bei den Lehramtsstudierenden hinsichtlich des Unterrichtsqualitätsmerkmals der Klassenführung unterschiedliche Profile der WIE extrahieren lassen

und inwiefern sich die extrahierten Profile in Bezug auf die von Schüler*innen wahrgenommene Klassenführung im Unterricht unterscheiden. Dafür wurde die WIE der Lehramtsstudierenden videobasiert mittels standardisierten Ratingitems sowie einer schriftlichen Analyse erfasst.

Beiträge des Symposiums

Von Wissen zu Performanz am Beispiel der Klassenführung im Sportunterricht (WiPe-Sport): Teilprojekt Wissenstest

Clemens Berthold, Matthias Baumgartner, Eric Jeisy
Pädagogische Hochschule St.Gallen

Theoretischer Hintergrund und Problemstellung

Das professionelle Wissen von Lehrpersonen gilt als integrativer Bestandteil aller Facetten professioneller Kompetenz (Blömeke et al., 2022) und hat einen positiven Einfluss auf das Niveau und die Stabilität der Un-terrichtsqualität (Voss et al.,2022). Diese Effektkette lässt sich im integrative Kompetenzmodell mit den drei Kompetenzfacetten (Blömeke et al., 2015) theoretisch begründen. Es gibt jedoch kaum empirische Er-gebnisse zu den vermuteten Wirkungs- und Entwicklungszusammenhängen der drei Kompetenzfacetten. Bisherige Untersuchungen fokussieren vor allem die Zusammenhänge zwischen jeweils zwei Facetten (ausg. Blömeke et al., 2022; Junker et al., 2021) und stützen sich dabei auf bereits validierte Instrumente (z.B. Junker et al., 2021; Voss et al, 2022), welche die einzelnen Facetten jedoch nicht immer trennscharf abbilden.

Im SNF-Forschungsprojekt WiPe-Sport sollen nun drei (facetten-)spezifische Instrumente zum Einsatz kommen, um die Klassenführung im Sportunterricht differenziert abzubilden. In diesem Beitrag werden die Generierung und die Validierung des Tests zur Erfassung des klassenführungsbezogenen Wissens vorge-stellt.

Methode

Neun Dimensionen der guten Klassenführung im Sportunterricht (Baumgartner et al., 2020) bildeten die Richtschnur für eine umfassende Literaturrecherche bezüglich des klassenführungsbezogenen Wissens, wobei der Fokus auf evidenzbasierte Praktiken gelegt wurde. Der erste Prototyp des Wissenstests sollte die vorhandenen (empirischen) Befunde möglich umfassend abbilden, wobei die unterschiedliche Qualität der Datenlage in den einzelnen Dimensionen deutlich wurde. So wurden wirksamkeitsorientierte und als auch normative Gütekriterien integriert. Zur Sicherung der inhaltlichen Validität des Wissenstests wurde mit diesem Prototyp eine dreistufige Delphi-Befragung durchgeführt, an der sieben Expert:innen aus dem Bereich der Sportdidaktik teilnahmen. Die überarbeiteten Items wurden anschließend im Rahmen einer Prä-Pilotierung angehenden sportunterrichtenden Lehrkräfte (N=54) vorgelegt und die Ergebnisse zusammen-geführt. Die inhaltlich validierten Items wurden in der folgenden Pilotierungsstudie von angehenden sport-unterrichtenden Lehrkräften aus vier Pädagogischen Hochschulen gelöst, wobei 877 Datensätze generiert werden konnten. Die Teilnehmenden haben per Zufall jeweils 1/3 aller Items gelöst. Anschliessend wurde mit Hilfe der Item-Response-Modellierung eine Bewertung der Items vorgenommen sowie eine Prüfung der internen Validität des Instruments durchgeführt. Über den Expertisevergleich von Studierenden unter-schiedlicher Studienjahre wurden Hinweise auf eine externe Validität geprüft.

Ergebnisse

Während der erste Prototyp 175 Items umfasste, führte der qualitative Prozess aufgrund von Sprache, Verständlichkeit oder Eindeutigkeit zu einer Reduktion auf 110 Items. Trotz hoher Lösungswahrscheinlich-keit in der Präpilotierung wurde eine große Anzahl an Items in die Pilotierung übernommen. Die statistische Analyse machte die unterschiedlichen Itemschwierigkeiten deutlich, was sich in einer besseren Passung des 2pl-Modells zeigte und den Ausschluss weiterer Items erforderte. Der vorliegende Test umfasst 78 Items, welche alle Dimensionen einer guten Klassenführung abdecken und keinen Misfit (0.8;1,2/-2;2) auf-weisen. Bei ausreichendem Modelfit ($\chi 2 = 17.985$, $p = 0.069$) ist anzunehmen, dass die Reliabilität (EAP = .603, WLE = .569) im Rahmen der Hauptuntersuchung ansteigt, wenn die einzelnen Studierenden die dop-pelte Anzahl an Items beantworten (Raju & Oshima, 2005).

Das vorliegende Testinstrument zielt insbesondere auf die Passung innerhalb des Forschungsprojekts ab und ermöglicht so eine inhaltlich stringente sowie statistisch valide Erfassung der Kompetenzfacette des professionellen (klassenführungsbezogenen) Wissens.

Von Wissen zu Performanz am Beispiel der Klassenführung im Sportunterricht (WiPe-Sport): Teilprojekt WIE-Test

Eric Jeisy, Matthias Baumgartner, Clemens Berthold
Pädagogische Hochschule St.Gallen

Theoretischer Hintergrund und Problemstellung

Die mediierende Kompetenzfacette der situativen kognitiven Wahrnehmungs-, Interpretations- und Ent-scheidungsfähigkeiten (WIE) von Lehrkräften verbindet das professionelle Wissen einer Lehrkraft mit ihrer unterrichtlichen Performanz und steht in einem positiven Zusammenhang mit den Lernleistungen von Schü-lerinnen und Schülern (Blömeke et. al., 2022). In Bezug auf die Klassenführung gilt als gesichert, dass sich die domänenspezifischen WIE-Fähigkeiten durch gezielte Schulung verbessern lassen (z.B. Seidel & Stür-mer, 2014) und es gibt Belege, dass Pädagogisches Wissen über Klassenführung mit den situationsspezi-fischen Fähigkeiten korreliert (z.B. König et al., 2014). Bezüglich des angenommenen Zusammenhangs zwischen den klassenführungsbezogenen WIE-Fähigkeiten und dem unterrichtlichen Handeln ist die Be-fundlage noch dünn und uneinheitlich (z.B. Gold et al., 2021; König & Kramer, 2016).

Für den Fachbereich Bewegung und Sport liegen zurzeit noch keine Studien vor, welche die Wirkungs- und Entwicklungszusammenhänge zwischen den verschiedenen Kompetenzfacetten im integrativen Kompe-tenzverständnis (Blömeke et al., 2015) systematisch untersuchen. Im Rahmen des vorzustellenden SNF-Projekts soll diese Forschungslücke am Beispiel der Klassenführung im Sportunterricht bearbeitet werden. Da bislang nur für den performativen Bereich der Klassenführung ein validiertes Testinstrument vorliegt (Baumgartner et al., 2000), wurden zur Erfassung der WIE-Fähigkeiten und des klassenführungsbezoge-nen Wissens je ein eigener Test entwickelt. In diesem Beitrag wird die Testentwicklung zur Erfassung der WIE-Fähigkeiten (angehender) Lehrkräfte im Fachbereich Bewegung und Sport vorgestellt.

Methode

Für den videobasierten Test wurden zehn Videovignetten zur Klassenführung im Sport entwickelt. Bei der Genese der Videovignetten wurde darauf geachtet, alle neun Basisdimensionen klassenführungsbezogener Performanz (Baumgartner et al., 2020) im Unterrichtsgeschehen abzubilden. Geschlossene Testfragen im Anschluss an die Betrachtung der Unterrichtssequenzen dienen dazu, die situationsbezogenen WIE-Fähigkeiten zu messen (vgl. Seidel & Stürmer, 2014). Zur inhaltlichen und strukturellen Test-Validierung wurden eine dreistufige Delphi-Befragung (Häder, 2014) mit acht Expert:innen der Sportdidaktik sowie eine Pretest-Erhebung mit Lehramts-Studierenden (N = 54) durchgeführt. Die empirische Konstruktvalidierung erfolgte mittels Pilotierungsstudie bei Studierenden aus vier Pädagogischen Hochschulen (N = 877). Dabei bearbeiteten die Proband:innen, jeweils zufallsbasiert, vier der zehn Vignetten. Die empirische Validität des Tests wurde intern mit Hilfe von Item-Response-Modellierung überprüft. Als externes Prüfkriterium wurde ein Vergleich der Testleistung der Studierenden in unterschiedlichen Studienphasen herangezogen.

Ergebnisse

Die qualitativen Rückmeldungen der Delphi-Befragung zeigten, dass alle neun Dimensionen guter Klassen-führung im Sport (Baumgartner et al., 2020) durch die Videovignetten repräsentiert werden und sich aus Sicht der Expert:innen als Item-Prompts zur Erfassung der WIE-Fähigkeiten von Sportlehrkräften eignen. Qualitative Überarbeitungen auf Item-Ebene im Anschluss an die Delphi-Befragung und den Pretest führten zu einer Reduktion der Itemanzahl (175 □ 160). Nach statistischer Analyse der Pilotierungsdaten erwies sich auch ein auf sieben Vignetten gekürztes Testinstrument als ausreichend reliabel (EAP = 0.647, WLE = 0.639). Die Wahl des 2pl – Modells bietet einen vergleichsweise besseren Modelfit und erlaubt auch inhalt-lich eine breitere Abbildung des Konstrukts. Es liegen keine Items mit auffälligen In bzw. Out-Fit-Indizes (0.8;1.2/-2;2) vor. Die unterschiedliche Verortung auf der Logit Skala von Items (M = -1.074 SD = 0.963) und Studierenden (M = 0.085, SD = 1.485) weist auf einen relativ einfachen Test hin. Die signifikant unter-schiedlichen Testwerte zwischen den Studierenden im ersten Studienjahr und dem letzten Studienjahr lie-fern einen externen Validitätshinweis.

Mit dem präsentierten WIE-Test liegt ein valides und reliables Testinstrument vor, mit dem in einer laufen-den Interventionsstudie die Teilfragen geklärt werden ob a.) die WIE (UV) ein bedeutsamer Prädiktor hin-sichtlich der klassenführungsbezogenen Performanzen (AV) von angehenden sportunterrichtenden Lehr-personen darstellen und b.) inwiefern diese Performanzen (AV) durch unterschiedliche Interventionspro-gramme (UV's) erhöht werden können. Die mithilfe dieses Testinstruments gewonnen Daten und Erkennt-nisse fliessen in die Gestaltung der klassenführungsbezogenen Ausbildung künftiger Lehrkräfte im Fachbe-reich Bewegung und Sport ein und sollen einen Beitrag zur Qualität des unterrichtlichen Bewegungs- und Sportangebots in den Schulen liefern.

Situierte Wahrnehmungs-, Interpretations- und Entscheidungsprozesse in Bezug auf Klassenführung – Eine Analyse latenter Profile von Lehramtsstudierenden

Madeleine Müller[1], Bernadette Gold[2]
[1]Universität Erfurt, [2]TU Dortmund

Theoretischer Hintergrund und Fragestellung

In den letzten Jahren stieg die Anzahl der Studien zu situierten Prozessen des Wahrnehmens, Interpretierens und Entscheidens (WIE) erheblich an (Amador et al., 2021, Santagata et al., 2021, König et al., 2022). WIE beschreibt die Fähigkeit, lernrelevante Situationen zu identifizieren, diese vor dem Hintergrund des professionellen Wissens zu interpretieren und davon abgeleitet eine passende Handlungsentscheidung zu treffen (Kaiser et al., 2015; Sherin & van Es, 2009). Zusammenhänge mit Unterrichtsqualität konnten vereinzelt gefunden werden (Gold et al., 2021, König & Kramer, 2016, König et al., 2021, Krauss et al., 2020), fokussieren jedoch eher einen variablenzentrierten Ansatz, welcher wenig Aufschluss darüber gibt, welches Fähigkeitsprofil von WIE günstig für effektiven Unterricht ist. Mittels eines personenzentrierten Ansatzes können intraindividuelle Unterschiede hinsichtlich WIE differenziert dargestellt werden und geben zudem Auskunft darüber, welches Profil sich als günstig für einen qualitätsvollen Unterricht herausstellt. Davon ausgehend werden in der vorgestellten Studie anhand des Unterrichtsqualitätsmerkmals Klassenführung folgende Fragestellungen fokussiert: Welche Profile der situierten Wahrnehmungs-, Interpretations- und Entscheidungsprozesse in Bezug auf Klassenführung können bei Lehramtsstudierenden identifiziert werden?

Methode

Die Stichprobe bestand aus 203 Masterstudierenden des Lehramtes (80.3% weiblich, MAlter=23.61 Jahre [SD=2.46 Jahre]). Die situierten WIE-Prozesse wurden videobasiert mittels standardisierter Ratingitems (adaptiert nach Gold & Holodynski, 2017) und einer schriftlichen Analyse mit Analyseprompts (adaptiert nach Gippert et al., 2022) erfasst. Die Ratingitems deckten das wissensbasierte Interpretieren von unterrichtsrelevanten Events ab, während die offenen Analysen in erkannte Events, deren wissensbasierte Interpretation und Generierung einer begründeten Handlungsalternative von zwei unabhängigen Rater*innen (ca. 30% doppelt; $\kappa_{erkennen}$=.74, $\kappa_{interpretieren}$=.74, $\kappa_{Handlungsalternative}$=.82) kodiert wurden. Von einer Teilstichprobe (n=74, 74.4% weiblich, MAlter=23.39 Jahre (SD=2.58) lagen zusätzlich Schüler*inneneinschätzungen zur wahrgenommenen Klassenführung aus Unterrichtsstunden während eines Praxissemesters vor. Auf Basis von z-transformierten Werten der Ratingitems und der offenen Analyse wurden mittels Two-Step-Clusteranalysen mit Log-Likelihood als Distanzmaß und dem Akaike-Informationskriterium (AIC) zunächst Profile identifiziert, um diese in der Teilstichprobe zu replizieren.

Ergebnisse

In der ersten Stichprobe konnten sechs Profile der situierten Wahrnehmungs-, Interpretations- und Entscheidungsprozesse unter den Lehramtsstudierenden identifiziert werden, wobei jeweils zwischen 24 bis 59 Lehramtsstudierende diese abdecken. Zwei Profile (P1, P2) zeigten jeweils hohe Ausprägungen in einem der beiden Erfassungsformate und zeigten demgegenüber unterdurchschnittliche Leistung bei den Ratingitems (P1: z=-1.78) oder bei der offenen Analyse (P2: -0.62<z<-0.31). Eine weitere Gruppe (P3) konnte eher einer analytischen Haltung zugeordnet werden und schnitt dementsprechend überdurchschnittlich im wissensbasierten Interpretieren (z=0.50) und den standardisierten Ratingitems (z=1.79) ab. Demgegenüber stand eine eher praktisch orientierte

Studierendengruppe (P4), welche zwar in den Ratingitems (z=0.18), im Erkennen (z=-0.18) sowie im Interpretieren (z=-0.36) eher durchschnittliche Leistungen erbrachte, jedoch überdurchschnittlich viele, begründete Handlungsalternativen (z=1.11) benennen konnte. Zuletzt wurden zwei Profile (P5, P6) identifiziert, welche entweder in allen Werten der WIE weit unter dem Durchschnitt lagen (P5: -1.29<z<-0.49) oder nur in der Lage waren, die Events besser zu erkennen (P6: z=0.93). Eine einfaktorielle MANOVA zeigte einen statistisch signifikanten Unterschied zwischen den Clusterzentren für die kombinierten abhängigen WIE-Variablen, $F(20, 638) = 50.32$, $p < .001$, $\eta p^2 = .54$, Wilk's $\Lambda = .043$. Post-hoc zeigte die für jede abhängige Variable durchgeführte einfaktorielle ANOVA, dass sich in jeder WIE-Variable die Cluster signifikant unterscheiden, was für die Validität der Cluster spricht. In der Teilstichprobe konnten vier vergleichbare Profile (P2, P3, P4, P5) identifiziert werden sowie ein „Mischprofil" (P1 & P6) mit jeweils 8 bis 19 zugeordneten Lehramtsstudierenden.

Bis zur Tagung werden die Unterrichtsqualitätsdaten der Schüler*innen ausgewertet sein und die Fragestellung präsentiert, inwiefern sich die identifizierten Profile in Hinblick auf ihre gezeigte Klassenführung im Unterricht unterscheiden. Neben Aufschluss über eine für qualitätsvollen Unterricht günstige Konstellation des WIE-Profils können die identifizierten Profile unter anderem auch für die Lehramtsausbildung hilfreich sein, um die Förderung der Teilprozesse der WIE differenziert und adaptiv anpassen zu können.

Fachspezifisches professionelles Wissen und Können von Sportlehrpersonen messen – Konstruktion und Validierung eines Testinstruments

Matthias Wittwer
Pädagogische Hochschule FHNW

Theoretischer Hintergrund

Als zentrale Faktoren professioneller Kompetenz gelten in der Professionsforschung das fachliche (CK) und das fachdidaktische (PCK) professionelle Wissen und Können von Lehrpersonen. Vor allem Letzterem wird auf Basis empirischer Studien eine zentrale Rolle im Hinblick auf die Qualität der Unterrichtshandlungen und die Lernleistungen von Schüler*innen zugeschrieben (z.B. Krauss et al., 2020). Als Folge ist auch im Fach Sport in jüngster Vergangenheit zunehmend das Forschungsinteresse zum Professionswissen von Lehrpersonen gewachsen und erste fachspezifische Instrumente zu dessen Erfassung wurden entwickelt (z.B. Heemsoth & Wibowo, 2020; Meier, 2020; Vogler et al., 2017). Diese klammern jedoch bislang den Bereich des CK grösstenteils aus. Zudem legen sie der analytischen kompetenzorientierten Forschungstradition folgend (vgl. Baumgartner, 2018) den Fokus primär auf explizites Wissen, währendem eine spezifische Erfassung impliziter Wissensanteile kaum untersucht wurde.

Fragestellung

Der vorliegende, im Rahmen der EPiC-PE 3:1-Studie (Messmer et al., 2022) zu verortende Beitrag geht somit der Frage nach, wie sich das fachspezifische professionelle Wissen und Können von Sportlehrpersonen empirisch strukturieren lässt. Neben einer primären Unterscheidung in CK und PCK wird dabei eine weitere auf Shulman (1986, 1987) zurückgehende Aufgliederung in Fakten- und Begründungswissen (bei CK) und Wissen bzw. Können bezüglich Instruktionsstrategien und (Miss-) Konzeptionen von Schüler*innen (bei PCK) angenommen. Zudem wird dem integrativen

Kompetenzstrukturmodell von Sportlehrpersonen (Baumgartner, 2018) folgend, sowohl dispositionales Wissen (analytische kompetenzorientierte Forschungstradition) als auch situationsspezifisches Wissen und Können (hybride kompetenzorientierte Forschungstradition) erfasst und damit eine Brücke zu einem «wisdom of practice» (Shulman, 1986) geschlagen. Hiermit wird ein Desiderat in der Professionsforschung angegangen, welches kontextualisierte Testverfahren fordert, die professionelle Kompetenz von Lehrpersonen nahe an der beruflichen Anforderung und damit ökologisch valide erfassen (König, 2015, Rutsch et al., 2018).

Methode

Im Rahmen einer Online-Testung wurden bei 181 angehenden bzw. unterrichtenden Sportlehrpersonen neben klassischen MC-Tests (51 Items) zur Erfassung dispositionalen Wissens auch sechs Video- und elf Textvignetten eingesetzt, um über die Darstellung von Unterrichtssituationen Wahrnehmungs-, Interpretations- und Entscheidungsprozesse und damit spezifische Wissensanwendungen erheben zu können. Während bei den Videovignetten (Fokus: CK) Bewegungsausführungen von Schüler*innen mit Optimierungspotenzial zu beurteilen sind (Wittwer, 2021), gilt es bei den Textvignetten (Fokus: PCK) in problembehafteten Unterrichtssituationen mögliche Handlungsalternativen bezüglich ihrer Güte auf einer Likertskala zu beurteilen. Die Auswertung geschah hierbei über Paarvergleiche der Urteile im Abgleich mit Expertennormen (Wittwer & Messmer, 2022). Inhaltlich fokussiert sich die Untersuchung in Anlehnung an das Kompetenzmodell von Messmer (2018) einerseits auf einen technisch-taktischen, andererseits auf einen technisch-gestalterischen Kompetenzbereich.

Ergebnisse und ihre Bedeutung

Die erhobenen Daten bestätigen die Zuverlässigkeit der verwendeten Testinstrumente und zeigen eine angemessene Passung zwischen Itemschwierigkeiten und Personenfähigkeiten (Wittwer & Messmer, 2022). Hinsichtlich der faktoriellen Struktur kann im Gegensatz zu den bisherigen Untersuchungen im Fach Sport die theoretische Annahme eines mehrdimensionalen fachspezifischen professionellen Wissens und Könnens jedoch nur teilweise unterstützt werden. Dies betrifft die beiden angenommenen Dimensionen CK und PCK, das dispositionale und situationsspezifische Wissen bzw. Können, die Subdimensionen innerhalb von CK (Fakten- und Begründungswissen) und auch die von Meier (2020) sowie Heemsoth und Wibowo (2020) nachgewiesenen Subdimensionen von PCK (Instruktionsstrategien & (Miss-)Konzeptionen von Schüler*innen). Mit Letzterem werden jedoch Ergebnisse in anderen Fächern bestätigt, bei denen keine empirische Differenzierung berichtet wurde (z.B. Lindl & Krauss, 2017). Nichtsdestotrotz stehen die Resultate damit durchaus im Einklang mit Shulmans (1987) Amalgamhypothese.

Das durch die Untersuchung gewonnene besseres Verständnis der strukturellen Grundlagen von Professionswissen im Fach Sport soll künftig eine Nachzeichnung von Professionalisierungsverläufen ermöglichen und somit Ausbildungsinstitutionen als Grundlage in der Lehre dienen. Zudem ermöglichen die reliablen und validen Daten zum fachspezifischen professionellen Wissen und Können von Sportlehrpersonen nun Wirkungsanalysen im Hinblick auf Unterrichtsqualität und Lernleistungen von Schüler*innen im Rahmen der EPiC-PE 3:1-Studie.

G-12: Mehr Studienmotivation in die Hochschulen: Wertüberzeugungen von Studierenden und ihre Förderung in spezifischen Studienkontexten

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* T03 R02 D26

Mehr Studienmotivation in die Hochschulen: Wertüberzeugungen von Studierenden und ihre Förderung in spezifischen Studienkontexten

Chair(s): **Lena S. Kegel** (Pädagogische Psychologie, Westfälische Wilhelms-Universität Münster)
*Diskutant*in(nen):* **Oliver Dickhäuser** (Pädagogische Psychologie, Universität Mannheim)

Studienmotivation gilt sowohl theoretisch als auch empirisch als ein Hauptfaktor für ein erfolgreiches Studium (Eccles & Wigfield, 2002; Schneider & Preckel, 2017). Allerdings berichten viele Studierende von motivationalen Problemen, die sogar als Studienabbruchgrund angegeben werden (Heublein et al., 2017). Gemäß der situativen Erwartungs-Wert-Theorie (Eccles & Wigfield, 2020) kommt Studienmotivation unter anderem durch einen hohen Wert des Studiums zustande. Wertüberzeugungen sind neben ihrer globalen Bedeutung für Studienmotivation auch für die Kurswahl und das Interesse an einzelnen Lehrveranstaltungen relevant (Harackiewicz & Priniski, 2018). Sie werden als multidimensionales Konstrukt mit den positiven Komponenten intrinsischer Wert (Freude), Wichtigkeit und Nützlichkeit (für persönliche Ziele, Beruf oder Studium) sowie der negativen Komponente Kosten verstanden (Anstrengungskosten, emotionale Kosten, Opportunitätskosten; Gaspard et al., 2015). Die Fragen, welche Wertüberzeugungen Studierende mit Blick auf ihr Studium haben und wie diese gefördert werden können, gewinnen in der empirischen Hochschulforschung zunehmend an Bedeutung.

Aktuelle Interventionsstudien adressieren mehrheitlich die Komponente der Nützlichkeit bei MINT-Studierenden, um Interesse und Studienleistungen zu fördern (Rosenzweig et al., 2022). Die erfolgreich evaluierten Interventionen beinhalten zumeist eine direkt kommunizierte Relevanz der Studieninhalte und/oder Reflexionsaufgaben zur Relevanz der Studieninhalte (Rosenzweig & Wigfield, 2016). Um die jeweils spezifischen Auswirkungen der Wertkomponenten auf das Studium zur Unterstützung von Studierenden zu nutzen, empfehlen Rosenzweig und Kolleg*innen (2022) Interventionen für weitere Wertüberzeugungen und Studienkontexte zu evaluieren.

Im vorliegenden Symposium werden Beiträge vorgestellt, die in verschiedenen Fachdisziplinen (Erziehungswissenschaft, Hochschuldidaktik, Pädagogische Psychologie) entstanden sind. Die Beiträge untersuchen Wertüberzeugungen sowie ihre Förderung in spezifischen Studienkontexten. Im ersten Beitrag werden auf globaler Ebene des Studiums die Wertüberzeugungen (intrinsischer Wert, Wichtigkeit, Nützlichkeit, Kosten) anhand einer groß angelegten Datenstichprobe querschnittlich auf Unterschiede zwischen neun Fächergruppen untersucht. Darauf aufbauend werden fachgruppenspezifische Zusammenhänge von Wertüberzeugungen mit Studiennoten und Studienabbruchintentionen aufgedeckt. Im Kontext des fachwissenschaftlichen Lehramtsstudiums analysiert der zweite Beitrag mithilfe von Daten aus einem randomisierten Laborexperiment studentische Sichtweisen und Wertüberzeugungen auf das Mathematikstudium im Spannungsfeld zwischen Fachwissenschaft und Berufsbezug. Der dritte Beitrag beleuchtet längsschnittlich die Wertüberzeugung Nützlichkeit und stellt eine Nützlichkeitsintervention zur Steigerung des Interesses an Forschungsmethoden im Vergleich zu einer behandelten Kontrollgruppe bei Studierenden der Erziehungswissenschaft unter Berücksichtigung der Leistungserwartung vor. Im vierten Beitrag wird

eine Intervention zur Reduktion motivationaler Kosten bei MINT-Studierenden präsentiert, die längsschnittlich als Feldstudie mit randomisierter behandelter Kontrollgruppe durchgeführt wurde. Zudem wird die Teilnahmebereitschaft an Interventionen und Evaluationsbefragungen untersucht. Abschließend wird die Bedeutung von Wertüberzeugungen für unser Verständnis und die Förderung der Motivation von Studierenden diskutiert. Dabei werden auch Grenzen des Ansatzes sowie Vor- und Nachteile fachgruppenspezifischer Analyseansätze beleuchtet.

Beiträge des Symposiums

"Warum lernen wir das?!" – Ein Vergleich subjektiver Wertüberzeugungen zwischen verschiedenen Studienfächern

Jonas Breetzke, Derya Özbagci, Carla Bohndick
Hamburger Zentrum für Universitäres Lehren und Lernen (HUL), Universität Hamburg

Wertüberzeugungen bezüglich des eigenen Studiums gelten als zentrale Prädiktoren für ein erfolgreiches Studium und gute akademische Leistungen (Eccles & Wigfield, 2020; Guo et al., 2016). Studierende, die ihrem Studium einen geringeren Wert zuschreiben, brechen es häufiger ab (Schiefele et al., 2007; Schnettler et al., 2020), während Studierende, die ihrem Studium einen hohen Wert beimessen, oft erfolgreicher in Ihrem Studienfach sind (Part et al., 2020; Wu et al., 2020).

In vielen Studienfächern nehmen diese positiven Wertüberzeugungen jedoch mit der Zeit ab (Kosovich et al., 2017; Robinson et al., 2019) und die meisten Studierenden stellen sich während seines Studiums einmal die Frage: "Warum lerne ich das?" (Hulleman & Harackiewicz, 2020). Die Antwort auf diese Frage ist prägend für Studienverläufe in allen Disziplinen. Doch obwohl disziplinäre Unterschiede ein wichtiges und immer wiederkehrendes Thema in der Hochschulforschung sind (Kuteeva & Airey, 2014; Ylijoki, 2000), fehlen Studien, die disziplinäre Unterschiede in den Wertüberzeugungen der Studierenden untersuchen.

Aufbauend auf der situativen Erwartungs-Wert-Theorie (Eccles & Wigfield, 2020), untersucht diese Studie deshalb anhand von neun Fächergruppen a) disziplinäre Unterschiede in subjektiven Wertüberzeugungen und b) den Zusammenhang zwischen Wertüberzeugungen und Studienerfolg für die verschiedenen Disziplinen.

Dazu wurde eine Datenstichprobe von N = 6844 Studierenden aus dem Nationalen Bildungspanel (NEPS) verwendet (Blossfeld & Rossbach, 2019; NEPS Network, 2021). Das Durchschnittsalter der Studierenden lag bei 29.63 Jahren (SD = 4.03) und 60.37 % der Teilnehmenden identifizierten sich als weiblich. Wertüberzeugungen wurden multidimensional durch den intrinsischen Wert, die berufliche Nützlichkeit, die Wichtigkeit und Kosten erhoben (Schiefele et al., 2002). Studienerfolg wurde durch die Durchschnittsnote sowie die Abbruchsintention (Trautwein et al., 2007) erfasst. Die Daten wurden mit Hilfe von Mehrgruppen Strukturgleichungsmodellen analysiert.

Die Ergebnisse unserer Studie zeigen, dass sich die Wertüberzeugungen der Studierenden stark zwischen den Studienfächern unterscheiden. Dabei werden verschiedene systematische Unterschiede deutlich: Zum Beispiel schätzen Studierende der Rechts- und Wirtschaftswissenschaften, als eher berufsbezogene Studiengänge (Schubarth & Speck, 2014), ihr Studium als deutlich nützlicher ein als ihre Kommilitonen in anderen Fächern es tun (βR = .270, p = .001 und βW = .389, p = .000). Studierende der Kunst- sowie der Sozialwissenschaften dagegen

zeigten deutlich höhere intrinsische Wertüberzeugungen (βK = .605, p = .000 und βS = .149, p = 001), sehen ihr Studium aber als weniger nützlich an als Studierende anderer Studiengänge (βK = -.213, p = .010 und βS = -.271, p = .000). Darüber hinaus zeigen unsere Ergebnisse, dass das Studienfach die Beziehung zwischen Wertüberzeugungen und Studienerfolg moderiert (Δ χ^2 (160) = 560.45, p = .000). Je nach Fachrichtungen profitieren die Studierenden demnach unterschiedlich stark von Ihren jeweiligen Wertüberzeugungen.

Insgesamt zeigen unsere Ergebnisse wie unterschiedlich die Wertüberzeugungen und deren Einfluss auf Studienerfolg in den jeweiligen Fachrichtungen sein kann. Auf praktischer Ebene kann dies dabei helfen motivationale Probleme der verschiedenen Fachrichtungen aufzudecken und zu minimieren. So haben zum Beispiel Studierende der Sozial- und Geisteswissenschaften das Gefühl, dass ihr Studium für ihre künftige Arbeit weniger nützlich ist, als ihre Kommiliton*innen in anderen Fächern es tun. Um hier Abhilfe zu schaffen, können Lehrende die berufliche Relevanz der Studieninhalte deutlicher hervorheben oder die im Studium erworbenen Fähigkeiten stärker betonen. Auch spezielle Interventionen (z.B. Nutzwertinterventionen; siehe auch Hulleman & Harackiewicz, 2020), die eine Verbindung zwischen den Inhalten des Studiums und dem (persönlichen und beruflichen) Leben der Studierenden herstellen, könnten in diesem Zusammenhang vielversprechend sein. In Anbetracht der deutlich geringen subjektiven Nützlichkeit der Sozial- und Geisteswissenschaften, scheint es durchaus realistisch, dass diese Studiengänge besonders von speziellen Interventionen profitieren könnten. Auf Grundlage dieser Ergebnisse erscheint es deshalb sinnvoll, die Wirksamkeit von Interventionen zur Steigerung von Wertüberzeugungen in zukünftigen Studien auch fachgruppenspezifisch zu untersuchen.

Zwischen Wissenschaft und Berufsbezug? Eine Untersuchung zu Sichtweisen von Lehramtsstudierenden auf die Fachwissenschaft Mathematik

Simon Guse, Dorothe Kienhues, Regina Jucks
Zentrum für Hochschullehre, Westfälische Wilhelms-Universität Münster

Im Studium entwickeln und reflektieren Studierende Wertüberzeugungen und Sichtweisen auf die studierte Fachwissenschaft. Dabei scheint Fachwissenschaft in der Lehre oft in einem Spannungsfeld zwischen einem wissenschaftlich-orientierten und einem berufsbezogenen Fokus zu stehen. Insbesondere im Lehramtsstudium wird dieses Spannungsfeld deutlich: Der für die Studienwahl maßgebliche Berufswunsch der Studierenden (Göller & Besser, 2021), scheint den rein fachwissenschaftlichen Anteilen im Studium gegenüber zu stehen. Wenn Verbindungen zwischen diesem fachlich-akademischen Hintergrund und der Mathematik als Vermittlungsgegenstand im Unterricht aufgezeigt werden, sehen Lehramtsstudierende mehr Nützlichkeit der Fachwissenschaft Mathematik (Koponen et al., 2016). Die berufsbezogene Relevanz ist nicht der einzige Grund für positive Wertüberzeugungen der Fachwissenschaft gegenüber, Lehramtsstudierende äußern auch Interesse an der Fachwissenschaft an sich (Lorentzen et al., 2019). Als Theorie-Rahmen zur Untersuchung dieser unterschiedlichen Sichtweisen auf die Fachwissenschaft Mathematik kann die Differenzierung von Wertüberzeugungen gemäß der Erwartungs-Wert-Theorie (Eccles & Wigfield, 2020) dienen. Wertüberzeugungen lassen sich differenzieren in intrinsischen Wert (Interesse und Freude), Wichtigkeit sowie Nützlichkeit (für kurzfristige oder langfristige Ziele). Eine übliche

kurzfristige Intervention zur Steigerung von Wertüberzeugungen ist das Adressieren von Relevanz des Inhalts oder Faches (Harackiewicz & Priniski, 2018).

Wir berichten Daten aus einer Studie (2x2-Design; N=259 Lehramtsstudierende Bachelor Mathematik) zur Untersuchung des Einflusses unterschiedlicher Relevanzbetonungen (mithilfe kurzer Zitate) auf Werteüberzeugungen zum fachwissenschaftlichen Studium. Es wurde manipuliert, ob die Zitate einen wissenschaftlichen Fokus (ja vs. nein) und einen berufsbezogenen Fokus (ja vs. nein) auf das Mathematikstudium beinhalten. Als abhängige Variablen wurden Wertüberzeugungen zum fachwissenschaftlichen Mathematikstudium mithilfe von sechsstufigen Likert-Skalen gemessen in den Facetten: intrinsischer Wert (Gaspard et al., 2015), persönliche Wichtigkeit (Gaspard et al., 2015), wahrgenommene Nützlichkeit für das Studium (Gaspard et al., 2015; Shin et al., 2019), wahrgenommene Nützlichkeit für das Unterrichten (Quast et al., 2021; Shin et al., 2019) und Freude an Mathematik (Meyer & Eley, 1999). Mithilfe einer offenen Frage wurde die persönliche Sichtweise auf die Fachwissenschaft Mathematik erfasst. Analysen zur Testung von Gruppenunterschieden zeigten keine signifikanten Unterschiede in den quantitativen abhängigen Variablen (MANOVA. Wissenschaftlicher Fokus: $F(5,251)=0.92$, $p=.487$; Berufsbezogener Fokus: $F(5,251)=0.95$, $p=.449$; Interaktion: $F(5,251)=1.51$, $p=.186$). Die Sichtweisen der Studierenden auf Mathematik scheinen eher stabil zu sein (Rach, 2022) und durch die gewählte niedrigschwellige Kurzzeitintervention nicht beeinflussbar. Da es keinerlei Unterschiede zwischen den Gruppen gab, konnten wir für nachfolgende Analysen den Datensatz als eine Stichprobe behandeln. Ausgehend von unserem Forschungsfrage nach den Werteüberzeugungen von Lehramtsstudierenden zur Fachwissenschaft Mathematik untersuchten wir mithilfe einer Clusteranalyse, ob sich die Lehramtsstudierenden verschiedenen Ähnlichkeitsgruppen zuordnen lassen.

Die Clusteranalyse (two-step, Silhouettenmaß 0.461) bezüglich der fünf Facetten von Wertüberzeugungen (intrinsischer Wert, persönliche Wichtigkeit, wahrgenommene Nützlichkeit für das Studium, wahrgenommene Nützlichkeit für das Unterrichten und Freude an Mathematik) als Inputvariablen schlägt drei disjunkte (alle $p<.001$) Profile vor. Im Profil „Affinität“ (26.3%) gaben die Teilnehmenden positive Wertüberzeugungen in allen fünf Facetten an ($4.40\leq M\leq 5.03$). Im Profil „Wertneutral“ (51.0%) wurden in allen Facetten Werte rund um die Skalenmitte angegeben ($3.30\leq M\leq 3.97$) und im Profil „Aversion“ (22.8%) niedrige Werte auf allen fünf Skalen ($1.93\leq M\leq 3.00$). Die qualitativen Daten unterstützen diese Profilbildung, beispielsweise gaben im Profil „Affinität“ signifikant mehr Teilnehmende (69.1%) eine positive wissenschaftlich-orientierte Sichtweisen auf Mathematik (z.B. Freude an Logik) an als im Profil „Wertneutral“ (35.6%, $p<.001$), gefolgt von denen im Profil „Aversion“ (5.1%, $p<.001$). Außerdem wurde eine negative berufsbezogene Sichtweise auf Fachmathematik wegen fehlender Relevanz signifikant seltener im Profil „Affinität“ (22.1%) als in den Profilen „Wertneutral“ (52.3%, $p<.001$) und „Aversion“ (66.1%, $p<.001$) genannt. Zwischen den drei Profile gab es keine Unterschiede in Bezug auf Alter, Fachsemester oder Zweitfach (alle $p>.084$) oder Experimentalbedingungen ($p=.306$).

Die Ergebnisse werden diskutiert in Bezug auf die Facetten der Sichtweisen auf die Fachwissenschaft (insbesondere Wissenschaftsorientierung und Berufsbezug). Zudem werden hochschuldidaktische Implikationen zum Umgang mit Studierenden aus den drei Profilen aufgezeigt.

Zur Wirkung einer Nutzwertintervention auf das Interesse von Studierenden der Erziehungswissenschaft an Forschungsmethoden

Laura Schäfer, Stefanie van Ophuysen
Institut für Erziehungswissenschaft, Westfälische Wilhelms-Universität Münster

Die forschungsmethodische Ausbildung von Studierenden stellt einen zentralen Bestandteil des Hochschulstudiums dar. Die zukünftigen Absolvent*innen sollen die Fähigkeit zum kritischen Denken sowie zum systematischen Problemlösen erlangen, um somit den komplexen Anforderungen des Arbeitsmarktes im 21. Jahrhundert zu genügen (Brew & Mantai, 2017).

Trotz der hohen Relevanz, die den Forschungsmethoden inner- aber auch außerhalb der Wissenschaft zugesprochen wird, sind Interesse und Leistung von Studierenden vornehmlich geistes- und sozialwissenschaftlicher Studiengänge – so des Studiengangs der Erziehungswissenschaft – diesbezüglich eher gering ausgeprägt (Slootmaeckers et al., 2014).

Mangelndes Interesse von Lernenden führt zu einer Verringerung der Lerneffektivität, was wiederum zu geringeren oder gar nicht genügenden Leistungen führt und im schlimmsten Fall der Ausschluss vom Hochschulstudium resultiert (Khavenson & Orel, 2014). Es ist also notwendig Studierende dazu zu motivieren sich engagiert mit Forschungsmethoden auseinanderzusetzen und ihre diesbezüglichen Kenntnisse auszubauen. Nicht nur um ihre Lerneffektivität und Leistung zu steigern oder ihnen den erfolgreichen Abschluss ihres Hochschulstudiums zu ermöglichen, sondern auch, um ihre Berufsfähigkeit und gesellschaftliche Teilhabe zu gewährleisten. Doch wie kann das Interesse und die Leistung der Studierenden gesteigert werden? Befunden der Motivationsforschung zufolge kann das Hervorheben des persönlichen Nutzens von Lerninhalten das Interesse und die Leistung von Lernenden steigern (z.B. Hulleman et al., 2010). Bei genauerer Betrachtung der Effekte dieser sogenannten „Nutzwertinterventionen“ zeigen sich jedoch uneindeutige Wirkungsunterschiede in Abhängigkeit der Fähigkeitsüberzeugungen der Studienteilnehmenden. So konnten unter anderem Canning und Harackiewicz (2015) zwar positive Effekte für Lernende mit hohem Vertrauen in die eigene Leistung feststellen, für Lernende mit geringem Vertrauen blieben diese Effekte jedoch aus. Dazu konträre Effekte fanden Hulleman et al. (2010): Die Betonung des Nutzens erwies sich als deutlich effektiver für Personen mit niedrigen Leistungserwartungen, im Vergleich mit Personen mit hohen Leistungserwartungen.

Auf Grundlage der bisherigen Erkenntnisse untersucht dieser Beitrag die Wirkung einer Nutzwertintervention auf das Interesse in Abhängigkeit der Leistungserwartungen. Die Nutzwertintervention wurde im Rahmen einer experimentellen Studie mit N = 34 Studierenden der Erziehungswissenschaft einer Universität mit dem Ziel der Steigerung des Interesses im Bereich der Forschungsmethoden durchgeführt. Die Teilnehmenden der Experimentalgruppe wurden über die Selbstgeneration des Nutzwerts dazu angeregt, die Relevanz von Forschungsmethoden für ihre persönliche oder berufliche Zukunft zu erkennen. Mittels Varianzanalysen wurde überprüft, inwiefern der wahrgenommene Nutzen unter Berücksichtigung der Leistungserwartungen die Entwicklung des Nutzwerts sowie des Interesses beeinflusst. Insgesamt können die Befunde den positiven Effekt der Nutzwertintervention auf den wahrgenommenen Nutzen und das Interesse nicht bestätigen. Über beide Gruppen hinweg ist ein generelles Absinken des wahrgenommenen Nutzens und des

Interesses über die Zeit festzustellen. Bezüglich der Leistungserwartungen zeigt sich jedoch ein signifikanter Interaktionseffekt für die Leistungserwartung, die Gruppe und den Messzeitpunkt des Interesses ($F(1,30) = 5.991$, $p = .02$, $\eta^2 = .166$). Während die Nutzwertintervention bei Studierenden mit geringen Leistungserwartungen einem Abfall des Interesses entgegenwirkt, wird das Absinken des Interesses bei Studierenden mit hohen Leistungserwartungen sogar unterstützt. Im Unterschied dazu zeigt sich in der Kontrollgruppe keine Veränderung des Interesses bei den Studierenden mit hohen, hingegen eine Abnahme des Interesses bei Studierenden mit geringen Leistungserwartungen. Die Befunde werden in den bisherigen Erkenntnisstand eingeordnet und mit Blick auf den Nutzen von Nutzwertinterventionen in diesem spezifischen akademischen Kontext diskutiert.

Motivationale Kosten im MINT-Studium reduzieren – Das Blended Learning-Programm UniProfi

Lena S. Kegel, Derya Turhan, Carola Grunschel
Pädagogische Psychologie, Westfälische Wilhelms-Universität Münster

Wie motiviert Studierende für ihr Studium sind, ist relevant für ihre Studienleistung (Schneider & Preckel, 2017) und ihren Verbleib im Studium (Heublein et al., 2017). Dabei wird Studienmotivation als multidimensionales Konstrukt verstanden, das sich nach der situierten Erwartungs-Wert-Theorie (SEVT; Eccles & Wigfield, 2020) aus einer Erwartungskomponente und differenzierten Wertkomponenten zusammensetzt (intrinsischer Wert, Wichtigkeit und Nützlichkeit als positive Wertkomponenten, Kosten als negative Wertkomponente). Da viele Studierende motivationale Herausforderungen berichten, sind Interventionen zur Förderung von Studienmotivation im Hochschulkontext nötig. Besonders MINT-Studierende könnten von Interventionsprogrammen profitieren, da in MINT-Studiengängen hohe Studienabbruchquoten auch auf eine geringe Studienmotivation zurückgeführt werden können (Heublein et al., 2022).

Interventionen zur Förderung von Studienmotivation fokussieren meist auf die Steigerung der Nützlichkeit des Studiums (Rosenzweig et al., 2022). Neuere Studien zeigen jedoch, dass die Förderung positiver Wertkomponenten nicht ausreicht, sondern besonders die motivationalen Kosten gesenkt werden müssen (Kim et al., 2021). Für die Reduktion von Kosten im Studium liegen zwei evaluierte Interventionen vor. Während Perez et al. (2019) keine Auswirkungen ihrer Intervention auf Kosten fanden, war die Intervention von Rosenzweig et al. (2020) in der Reduktion der Kosten erfolgreich - allerdings nur für Studierende mit geringeren Studienleistungen. In der Hochschulpraxis nehmen jedoch besonders leistungsstarke Studierende an Unterstützungsangeboten teil (Falk & Marschall, 2021), sodass Studierende mit potentiell größerem Bedarf nicht erreicht werden. Zudem fehlte in beiden Interventionsstudien eine differenzierte Betrachtung der Kosten, obwohl eine Differenzierung von Anstrengungskosten, emotionalen Kosten und Opportunitätskosten für die Erklärung von Studienerfolg aufschlussreich ist (Flake et al., 2015).

In der vorliegenden Studie wird ein Interventionsprogramm zur Reduktion von Anstrengungskosten, emotionalen Kosten und Opportunitätskosten in MINT-Studiengängen evaluiert. Zuerst wird untersucht, ob Studierende mit vollständiger Teilnahme an der Interventionsstudie vor der Intervention adaptiveres Studienerleben und Lernverhalten berichten als Studierende mit unvollständiger Teilnahme. Außerdem wird überprüft, ob Studierende der Interventionsgruppe nach der Teilnahme geringere Kosten sowie höhere Erwartungen, positive Wertausprägungen und Studienzufriedenheit

berichten als Studierende der Kontrollgruppe. Zudem wird angenommen, dass die Reduktion der Kosten für leistungsschwächere Studierende höher ausfällt als für andere Studierende.

Die Evaluation des Interventionsprogramms wurde als randomisiertes Kontrollgruppendesign im Sommersemester 2022 durchgeführt. Physik- und Mathematikstudierende des zweiten Fachsemesters einer Universität wurden eingeladen, am Interventionsprogramm zur Reduktion von Kosten oder am Kontrollprogramm zur Förderung von Kommunikationskompetenzen teilzunehmen. Beide Programme wurden im Blended Learning-Format über neun Wochen umgesetzt. In jeder Programmwoche bearbeiteten die Teilnehmenden in 30 Minuten spezifisch für MINT-Studierende aufbereitete Videos, Reflexionsaufgaben sowie Übungen mit eigenen Beispielen. Die Befragungen zu Studienmotivation, Lernverhalten, Studienzufriedenheit und Leistungen fanden vor Interventionsstart (Prätestung, n = 262), nach Interventionsabschluss (Posttest, n = 108) und zu Beginn des Wintersemesters 2022/2023 (Follow-Up-Test, ausstehend) statt. Als vollständige Teilnahmen wurden für die Analysen n = 55 Studierende definiert, die an mindestens fünf Interventionswochen, dem Prätest und dem Posttest teilgenommen haben.

t-Tests für unabhängige Stichproben zeigten adaptivere Ausprägungen im Prätest von Abiturnote, Opportunitätskosten, investierter Anstrengung und Studienzufriedenheit für Studierende mit vollständiger Teilnahme im Vergleich zu Studierenden mit unvollständiger Teilnahme. Die Wirksamkeit des Interventionsprogramms in Bezug auf die Studienmotivation und Studienzufriedenheit wurde mit Varianzanalysen mit Messwiederholung (2x2-faktoriell) anhand der Studierenden mit vollständiger Teilnahme analysiert, wobei keine Unterschiede zwischen der Interventionsgruppe und der Kontrollgruppe festgestellt wurden. Zusätzlich zeigten Moderationsanalysen, dass leistungsschwächere Studierende nicht stärker vom Interventionsprogramm profitierten als andere Studierende. Langfristige Effekte des Interventionsprogramms werden anhand der Follow-Up-Testung noch analysiert.

Insgesamt nahmen eher leistungsstarke MINT-Studierende mit weniger motivationalen Herausforderungen an dem Interventionsprogramm teil. Wie in früheren Studien (Perez et al., 2019; Rosenzweig et al., 2020) bewirkte das Interventionsprogramm direkt nach Abschluss der Intervention im Vergleich zu einer behandelten Kontrollgruppe keine Reduktion von Kosten. Für zukünftige Studien sollte vor allem eine Erhöhung der Teilnahme von leistungsschwächeren Studierenden angestrebt werden, indem Interventionen durch Fächer gut unterstützt in das Studium integriert werden.

G-13: Wirksamkeit und Transferpotentiale von Berufsorientierungsmaßnahmen

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* T03 R02 D39

Wirksamkeit und Transferpotentiale von Berufsorientierungsmaßnahmen

Chair(s): **Birgit Ziegler** (Technische Universität Darmstadt, Germany), **Katja Driesel-Lange** (Wstfälische Wilhelms Universität Münster), **Mona Granato** (Bundesinstitut für Berufsbildung, Bonn)
*Diskutant*in(nen):* **Elena Makarova** (Universität Basel)

Jugendliche und Erwachsene in der Gestaltung ihrer (Aus)Bildungs- und Erwerbslaufbahn zu begleiten und zu unterstützen, ist in modernen, hoch arbeitsteilig strukturierten Gesellschaften eine wichtige bildungspolitische Herausforderung. „Berufswahl“ beginnt in der Kindheit, ihre Konkretisierung im ersten Übergang in Ausbildung oder Studium ist eine zentrale Entwicklungsaufgabe des Jugendalters, die den Anfang einer dynamischen, kontinuierlich veränderbaren Laufbahn markiert.

Berufliche Aspirationen werden schon früh durch Geschlecht und soziales Prestige beeinflusst und Berufe werden entlang stereotyper Vorstellungen aus dem individuellen Aspirationsfeld ausgeschlossen. Am Beginn schulischer Berufsorientierung sind Jugendliche auf einem individuell je unterschiedlichen Entwicklungsstand hinsichtlich berufswahlrelevanter Kompetenzen, Ziele und Aktivitäten. Ressourcenausstattung und Unterstützungsaktivitäten im familialen Umfeld prägen berufliche Übergänge. Jugendlichen aus sozialen Risikolagen sind hierbei seltener erfolgreich. Zudem wird das Übergangsgeschehen von strukturellen und konjunkturellen Faktoren beeinflusst. Jugendliche benötigen Unterstützung in der Berufsorientierung, was sich in den Quoten von Ausbildungs- und Studienabbrüchen sowie in wachsenden Passungsproblemen auf dem Ausbildungsmarkt offenbart. Der Rückgang von Berufswahlaktivitäten während der Pandemie unterstreicht die Relevanz von BO-Maßnahmen.

Berufsorientierungsmaßnahmen können unterschiedliche Ziele verfolgen. Die Förderung von Berufswahlkompetenz steht im Mittelpunkt einer erfolgreichen Berufsorientierung. Darüber hinaus kann die Erweiterung des Berufswahlspektrums als zentrale Aufgabe von Berufsorientierung verstanden werden, um z. B. Geschlechtersegregation und sozialen Disparitäten bei der Berufswahl entgegenzuwirken und dadurch Passungsprobleme auf dem Ausbildungsmarkt bzw. den Mangel an Fachkräften in bestimmten Berufsfeldern (z. B. MINT, Pflege) zu reduzieren. Abhängig von den Zielen und von berufswahltheoretischen Prämissen werden unterschiedliche Berufsorientierungsmaßnahmen an Schulen implementiert. Die Befundlage zu ihrer Wirksamkeit gilt weiterhin als eingeschränkt.

Wenngleich sich die Berufsorientierungsforschung hierzulande in den letzten Jahren intensiviert hat, ist man von einer evidenzbasierten Gestaltung von Berufsorientierung an Schulen noch weit entfernt. Das geplante Symposium greift diese Forschungslücke auf und stellt empirische Befunde zu Berufsorientierungsangeboten vor. Es werden Erkenntnisse aus Interventionsstudien und zur Gestaltung von Berufsorientierungsmaßnahmen zur Diskussion gestellt. Der erste Vortrag präsentiert Befunde zum Einfluss von „Ausbildungsbotschafter:innen“. Jugendlichen werden durch positive Role Models verschiedene Ausbildungsberufe nähergebracht. Die Ausbildungsbotschafter vermitteln nicht nur Informationen, sondern zeigen auch Identifikationspotentiale auf. Der zweite Vortrag stellt Befunde zum Workshop „Logiken der Berufswahl“ vor. Dieses Konzept adressiert explizit das soziale

Anerkennungsbedürfnis in der Berufsorientierung. Jugendliche werden zum Nachdenken angeregt, inwieweit sie ihre Berufswünsche am sozialen Anerkennungspotential ausrichten. Der dritte Vortrag widmet sich der Sicht von Eltern. Präsentiert werden Befunde über die Zufriedenheit der Eltern mit der Berufswahl ihrer Kinder. Das Einverständnis der Eltern stellt sich als erklärungsmächtig für die Unterstützungsleistung in Gesprächen über die berufliche Zukunft heraus.

Beiträge des Symposiums

„Ich wär' so gern wie du" – Können Rollenvorbilder die beruflichen Aspirationen junger Menschen verändern? Ergebnisse einer Interventionsstudie

Janina Beckmann[1], **Mona Granato**[2]
[1]Bundesinstitut für Berufsbildung, [2]Bundesinstitut für Berufsbildun

Fragestellung

Immer mehr junge Menschen bevorzugen akademische Bildungswege, statt einer dualen Ausbildung. Gerade handwerkliche und auch gewerblich-technische Ausbildungsbildungsberufe sind mehrheitlich durch eine geringere Nachfrage gekennzeichnet, wodurch sich der Fachkräftemangel in Deutschland weiter verschärft (BIBB-Datenreport, 2022). Die Steigerung der Attraktivität der beruflichen Ausbildung ist daher ein zentrales bildungspolitisches Anliegen.

Ein neuartiger Ansatz, das Interesse an Ausbildungsberufen zu erhöhen, stellt das Berufsorientierungsangebot „Ausbildungsbotschafter" der IHK und HWK dar. Anhand beruflicher Rollenvorbilder soll die Attraktivität der dualen Ausbildung bei Jugendlichen gesteigert werden. Dabei stellen Auszubildende „auf Augenhöhe" ihren Beruf Schüler*innen vor und beantworten Fragen zu ihrem Ausbildungs- und Berufsalltag. Sie vermitteln nicht nur Wissen über ihren Beruf, sondern werden zu Rollenvorbildern, mit denen sich Jugendliche identifizieren können (Athanasiadi u.a. 2020).

Dieser Beitrag untersucht am Beispiel des Berufsorientierungsangebots „Ausbildungsbotschafter", inwieweit berufliche Rollenvorbilder berufliche Aspirationen bei Schüler*innen steigern können. Im Zentrum stehen folgende Fragen: (i) können berufliche Rollenvorbilder das Interesse, eine berufliche Ausbildung aufzunehmen, erhöhen? und (ii) können berufliche Rollenvorbilder die beruflichen Aspirationen für bestimmte Ausbildungsberufe (handwerklich-technischer sowie kaufmännischer Bereich) verändern?

Theorie

Soziale Einflüsse spielen eine wichtige Rolle in der Berufswahl junger Menschen. Gut belegt ist, dass die Wahl eines Berufs nicht allein auf Basis der eigenen Fähigkeiten und Interessen erfolgt, sondern durch soziale Einflüsse mitbestimmt wird (z.B. Matthes u.a. 2019). So sind berufliche Aspirationen wesentlich durch wahrgenommene Genderrollen und das soziale Prestige eingegrenzt (Gottfredson 1996). Signifikante Andere wie Eltern und Freunde begleiten den Berufswahlprozess junger Menschen durch längerfristige Sozialisations- und Interaktionsprozesse. Aber auch kurzfristige soziale Einflüsse können für die Herausbildung beruflicher Aspirationen bedeutsam sein. Die „Ausbildungsbotschafter" stellen solche berufliche Rollenvorbilder dar, die sich durch eine einmalige und kurzfristige Begegnung auszeichnen.

Berufliche Rollenvorbilder können durch unterschiedliche Mechanismen für die Veränderung beruflicher Aspirationen bedeutsam werden. Dabei wirken berufliche Rollenvorbilder wesentlich durch motivationale Prozesse (Morgenroth u.a. 2015). Demnach können positive und erstrebenswerte Rollenvorbilder als Inspiration dienen und die (möglicherweise stereotype oder negative) Wahrnehmung einer beruflichen Ausbildung oder eines Berufs in eine positive Richtung verändern. Berufliche Rollenvorbilder verkörpern außerdem „Beispiele des Möglichen“ und können so das Vertrauen in die eigenen berufsspezifischen Fähigkeiten erhöhen (Bandura 1997). Darüber hinaus stellen Rollenvorbilder berufsspezifische Informationen bereit, u.a. über die beruflichen Tätigkeiten und Arbeitsbedingungen und können dadurch die Kenntnis über verschiedene Ausbildungsberufe vermitteln.

Empirische Studien konnten zeigen, dass berufliche Rollenvorbilder z.B. die beruflichen Aspirationen verändern können. Der Fokus liegt jedoch bislang auf (akademischen) Rollenvorbildern aus MINT-Berufen (z.B. Lawner u.a. 2019; van Camp u.a. 2019). Inwieweit Rollenvorbilder aus nicht-akademischen Berufen wirksam sind, ist bislang kaum erforscht. In dieser Studie soll daher auch die Frage beantwortet werden, inwieweit berufliche Rollenvorbilder aus kaufmännisch und handwerk-lich-technischen Ausbildungsberufen berufliche Aspirationen junger Menschen verändern. Dabei wird auch in den Blick genommen, welche Merkmale des Berufsinhabers (z.B. vorgestellter Beruf, Identifikation) die Wirksamkeit erhöhen.

Methode

Die Wirksamkeit von Rollenvorbildern wird anhand einer quasi-experimentellen Studie untersucht, in welcher zwei Gruppen von Schüler*innen zu drei Zeitpunk-ten zu ihren (beruflichen) Aspirationen befragt werden. Die Treatmentgruppe nimmt an den Vorträ-gen der „Ausbildungsbotschafter“ teil, die Kontrollgruppe hingegen erst nach dem Befragungszeit-raum. Ergebnisse zur Wirksamkeit von Rollenvorbildern werden mittels Mittelwertvergleichen und Regressionsverfahren auf Basis der Vorher-Nachher-Befragung untersucht.

Ergebnisse

Erste Ergebnisse zeigen im Hinblick auf berufliche Aspirationen, dass die Teilnahme an den Ausbildungsbotschaftervorträgen die berufliche Aspiration für den vorgestellten Ausbildungsberuf im Durchschnitt leicht erhöhen kann. Die Wirksamkeit variiert dabei stark nach vorgestelltem Ausbildungsberuf und der wahrgenommenen Identifikation mit dem Rollenvorbild. Insgesamt zeigt sich, dass berufliche Rollenvorbilder dazu beitragen können, das Interesse an Ausbildungsberufen zu erhöhen. Die weiteren Analysen fokussieren auf die erste Forschungsfrage (Erhöhung des Interesses an Berufsausbildung) sowie auf die Bedeutung des Geschlechts für die Wirksamkeit von Rollenvorbildern, d.h. darauf, wie sich Rollenvorbilder aus einem männlich dominierten Ausbildungsberuf auf die Aspi-rationen differenziert nach Geschlecht auswirken.

„Logiken der Berufswahl“– Wirkung und Einflussfaktoren der Auseinandersetzung mit dem Be-dürfnis nach sozialer Anerkennung in der Berufswahl

Sevil Mutlu[1], Stephanie Oeynhausen[2]
[1]Technische Universität Darmstadt, [2]Bundesinstitut für Berufsbildung

Trotz der Vielfalt an Berufsorientierungsmaßnahmen, die Jugendlichen als Unterstützung im Berufswahlprozess angeboten werden, ist die empirische Evidenz zur Wirksamkeit von Berufsorientierungs-

angeboten bislang begrenzt (Brüggemann et al., 2017; Hirschi & Froidevaux, 2020). Ausgehend vom Konzept einer „anerkennungssensiblen“ Berufsorientierung wurde eine Interventionsstudie durchgeführt, die hier ansetzt und zur Weiterentwicklung einer evidenzbasierten BO-Praxis beitragen soll.

Theoretischer Hintergrund und Fragestellung

Anerkennungssensible Berufsorientierung unterstellt, dass Jugendliche von Berufen attrahiert sind, die ein hohes Identifikationspotential bieten (Oeynhausen & Ullrich, 2020), und dass die Antizipation mangelnder Anerkennung eines Berufs im sozialen Umfeld ein zentraler Ausschlussgrund (Aversionsfaktor) von Berufen ist (Matthes, 2019). Weiterhin wird ausgehend von der Eingrenzungs- und Kompromisstheorie (L.S. Gottfredson, 1981: passim) sowie auf Basis empirischer Evidenz davon ausgegan-gen, dass sich berufliche Aspirationen einer Ausschlusslogik folgend entwickeln und Jugendliche am Ende ihrer Schulzeit bereits ein stark eingegrenztes berufliches Aspirationsfeld in den Dimensionen Geschlechts- und Prestigepassung aufweisen (Helwig, 2001; Ratschinski, 2009; Steinritz et al., 2016; Kleinert & Schels, 2020), und dass berufliche Aspirationen berufliche Entscheidungsprozesse nach-weislich auf lange Sicht beeinflussen (Rojewski, 2005; Düggeli & Neuenschwander, 2015; Lawson et al., 2018). Die Reflexion beruflicher Aspirationen wird daher zum Ausgangspunkt einer anerkennungs-sensiblen Berufsorientierung.

Ausgehend vom skizzierten Theorierahmen wurde der Workshop „Logiken der Berufswahl“ als Intervention entwickelt. Jugendliche werden anregt und ermutigt, ihr Bedürfnis nach sozialer Anerkennung in ein reflektiertes Verhältnis zu anderen Bedürfnissen zu setzen. Spielerische Werkzeuge und Metho-den laden die Teilnehmenden ein, ihre eigenen Geschlechter- und Prestigestereotypen zu diskutieren und dabei zu reflektieren, wie diese ihre beruflichen Aspirationen beeinflussen (Oeynhausen & Mutlu 2022). Die Intervention zielt darauf ab, weitgehend unbewusstes Verhalten beim Ausschluss von Beru-fen zu unterbrechen und einen eher "reflexiv-kalkulierenden" Modus bei der Berufswahl (Esser, 2005) zu unterstützen.

Eine zentrale Fragestellung ist, ob sich über diese Form der Intervention – also über einen expliziten Zugang zur Thematik – die Bereitschaft und Art der Jugendlichen verändert, sich mit der eigenen Berufswahl auseinanderzusetzen und ob dadurch ein höheres Bewusstsein im Hinblick auf die soziale Diskreditierung von Berufen entsteht und dies gegebenenfalls einen Einfluss auf die Einschätzung von Berufen hat. Diese Auseinandersetzung kann sich unter anderem auf das Interesse am Thema Berufs-wahl auswirken, sowie auf motivationale Aspekte, wie die Reflexion über die eigenen Bedürfnisse im Kontext der Berufswahl als auch auf weitere Handlungsintentionen im Berufsorientierungsprozess, die das Ziel einer selbstgesteuerten Berufswahl verfolgen.

Methode

Die Wirksamkeit des Workshops wird in einem quasi-experimentellen Design in zwei Regionen mit bis zu drei Messzeitpunkten (M1-M3) unter-sucht. Die Interventionsgruppe wird vor und nach der Intervention (M1/M2), sowie vier bis sechs Wo-chen (M3) später befragt. Die Wartekontrollgruppe wird zu M1 und M3 befragt. Zur Grundgesamtheit zählen Schüler:innen der Jahrgangsstufen 8 bis 11 allgemeinbildender Schulen. Die erreichte Stich-probe umfasst n = 688 Treatmentschüler:innen und n = 337 Kontrollschüler: innen. Es erfolgt zunächst ein globaler Vergleich zu abhängigen Variablen auf

Basis der M1 und M3-Daten über T-Tests. Im An-schluss werden differenzierte Analysen zu Einflussfaktoren, wie beispielsweise die Bewertung des Workshops zu M2, antizipierte Erwartungen der Eltern und Kontextfaktoren wie Schulart und Ge-schlecht auf Basis von Regressionen und MANOVAS durchgeführt.

Ergebnisse

Die Datenerhebung ist seit Ende September abgeschlossen. Neben der Konzeption des Workshops werden im Vortrag erste Analysen zur der M1-M3 Messung in Bezug auf die abhängigen Variablen „Interesse am Thema Berufswahl“ und die „Entscheidungssicherheit im Berufswahlprozess“ sowie Ergebnisse differenzierter Analysen zum Einfluss abhängiger Variablen präsentiert. Im Mittelpunkt dieser ersten Analysen steht die Frage, inwieweit sich durch die Teilnahme an der Intervention (1) das Interesse der Schüler:innen am Thema Berufswahl erhöht (2) sich ihre Entscheidungssicherheit im Berufswahlprozess verändert hat.

Die Sicht der Eltern auf die Berufswünsche Ihrer Kinder: Welche Rolle spielt die Zufriedenheit der Eltern mit der Berufswahl ihrer Kinder für Eltern-Kind-Gespräche?

Katja Driesel-Lange[1], Jerusha Klein[1], Lore Funk[2], Wenka Wenzel[2]
[1]Westfälische Wilhelm-Universität Münster, [2]Kompetenzzentrum Technik-Diversity-Chancengleichheit e. V. Bielefeld

Die Entwicklung einer beruflichen Perspektive und daran anknüpfend die Gestaltung eines nachschulischen Bildungswegs ist eine zentrale Entwicklungsaufgabe im Jugendalter, die neben anderen Entwicklungsaufgaben zeitgleich bearbeitet werden muss (Eschenbeck & Knauf, 2018). In ihrem Verlauf werden Heranwachsende mit spezifischen berufswahlbezogenen Aufgaben konfrontiert, um zu einer tragfähigen Entscheidung zu gelangen (Bergmann, 2004). Um die Entwicklungsaufgabe der Berufswahl zu meistern, benötigen Jugendliche Unterstützung. Diese wird zum einen in Form flankierender pädagogischer Angebote der schulischen Berufsorientierung gewährleistet. Zum anderen profitieren Jugendliche besonders von sozialer Unterstützung, die u.a. Lehrpersonen und Eltern bereitstellen (Schindler, 2012, Mayhack, 2011).

Eltern sind in zweierlei Hinsicht bedeutsam. Zum einen repräsentieren sie einen starken exogenen Einflussfaktor auf Bildungs- und Berufsverläufe (Kracke, Middendorf & Buck, 2018) und gelten damit mittelbar als Gatekeeper für berufliche Entscheidungsprozesse. Zum anderen nehmen sie durch ihren Erziehungsstil und ihr Engagement Einfluss auf berufliche Entwicklungsprozesse (Neuenschwander, 2008; Kracke & Noack, 2005). Sie ergreifen vor diesem Hintergrund auch die Rolle als informelle Berufsberater und werden für das Vorankommen im Berufswahlprozess von den Jugendlichen selbst unterstützender wahrgenommen als Lehrpersonen und professionelle Berufsberater*innen (Ulrich, Frey, Ertelt & Ruppert, 2019). Die Unterstützung, die Eltern leisten, kann informationeller, instrumenteller oder emotionaler Natur sein. So weisen sie auf Veranstaltungen hin, organisieren Praktika oder geben Halt bei Rückschlägen (Ulrich et al., 2019; Mayhack, 2011). Eine wichtige Komponente der Unterstützung bilden Gespräche (Dietrich & Kracke, 2009; Gehrau, 2019). Bisher wurden Gesprächsinhalte jedoch wenig differenziert in Studien beleuchtet bzw. vor allem aus der Sicht der Jugendlichen betrachtet. Vor diesem Hintergrund nimmt die zu präsentierende Studie die Perspektive der Eltern auf und verfolgt die Frage, welche Gesprächsinhalte für Sorgeberechtigte im

Kontext der Berufswahl relevant sind, wie häufig Gespräche geführt werden und ob Einflüsse der Bildungsherkunft, des Geschlechts und der Schultypen zum Tragen kommen.

Zur Beantwortung der Fragestellung wurden Eltern mit Kindern im Alter von 10 bis 20 Jahren auf mehreren Wegen zur Teilnahme an einer Online-Befragung eingeladen. Ein Rekrutierungsansatz war, Eltern über das Anmeldeverfahren zu den Aktionstagen „Girls' Day" und „Boys' Day" zu gewinnen. Weiterhin unterstützten der Bundeselternrat und die Landeselternräte, indem sie Eltern über ihre Verteiler zur Befragung informierten und den Link zum Fragebogen verbreiteten. Auch konnten Teilnehmende den Fragebogen-Link nach dem Schneeballprinzip an potenzielle Interessenten weiterleiten. Insgesamt haben bundesweit 3.657 Zielpersonen an der Befragung teilgenommen, 1.599 vollständig ausgefüllte Fragebögen konnten für diesen Beitrag in die Analysen einbezogen werden. Mit 12 Items wurden die Inhalte und die Häufigkeit der Gespräche auf einer 7-stufigen Likertskala (Cronbachs Alpha = .92) erfasst. Grundlage bildet eine neu entwickelte Gesprächsskala, die anschlussfähig ist an kommunikationswissenschaftliche Forschungsarbeiten zur Berufsorientierung (Gehrau, 2019). Neben einer deskriptiven Betrachtung, um die Kontexte der wahrgenommenen Gespräche herauszuarbeiten, wurden zudem über eine Regression verschiedene unabhängige Variablen (Bildungsherkunft, Geschlecht Eltern/Kind, Alter Kind, Zufriedenheit mit Berufswunsch) als Einflussfaktoren getestet.

Insgesamt zeigt sich in der vorliegenden Stichprobe, dass Eltern zu einem großen Maß im Gespräch mit ihren Kindern über die Berufswahl sind. Differenzierte Analysen machen Schulformunterschiede sichtbar: Mit Jugendlichen an Hauptschulen finden häufiger Gespräche mit den Eltern statt. Darüber hinaus wird deutlich, dass in der Gegenüberstellung der Gesprächskontexte Probleme und Ängste bei der Berufswahl weniger häufig thematisiert werden. Gesprächsinhalte, die auf eher informationelle und instrumentelle Unterstützung abzielen, finden dagegen häufiger statt. Bemerkenswert ist, dass sich in der Datenanalyse die Zufriedenheit über den Berufswunsch des Kindes als Prädiktor für die Häufigkeit der Gespräche zeigt. Zufriedenere Eltern sprechen häufiger mit ihren Kindern.

Die Ergebnisse der Studie sind zum einen in den Diskurs um die Bedeutung der Eltern für die Förderung individueller beruflicher Entwicklung im Jugendalter und auch in Bezug auf mögliche potenzielle Problemlagen aufzunehmen. Zum anderen soll das methodische Vorgehen reflektiert und vor dem Hintergrund innovativer Ansätze der Datenauswertung diskutiert werden.

G-14: Stereotype von (angehenden) Lehrkräften

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* T03 R02 D81

Stereotype von (angehenden) Lehrkräften

Chair(s): **Oscar Yendell** (Universität Mannheim), **Hanna Beißert** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation)
*Diskutant*in(nen):* **Anke Heyder** (Ruhr-Universität Bochum)

Stereotype – verstanden als generalisierte Überzeugungen über Angehörige verschiedener sozialer Gruppen (Eagly & Chaiken, 1993) – zeigen sich bei angehenden und praktizierenden Lehrer*innen (Martiny & Froehlich, 2020). Diese lassen sich sowohl auf der individuellen Ebene als auch auf der kulturellen Ebene erfassen, die umschreibt, welche Stereotype innerhalb einer Gesellschaft geteilt werden (Cuddy et al., 2007; Eagly & Mladinic, 1989). Stereotype umfassen unterschiedliche Inhaltsbereiche und Valenzausprägungen (Bonefeld & Karst, 2020; Fiske et al., 2002) und können zu einer kategorienbasierten und vereinfachten Informationsverarbeitung beitragen, die auf der sozialen Zugehörigkeit der jeweilig stereotypisierten Person basiert (Fiske et al., 2018).

Im Schulkontext wurde bereits die Existenz verschiedener Stereotype nachgewiesen, wie beispielsweise bezüglich des sozioökonomischen Status (Dunkake & Schuchart, 2015; Glock & Kleen, 2020), des Migrationshintergrundes (Froehlich et al., 2016), des Geschlechtes (Muntoni & Retelsdorf, 2020) oder sonderpädagogischer Förderbedarfe (Bennett et al., 2018; Pit-ten Cate & Krischler, 2020). Stereotype können mit Kausalattributionen zusammenhängen und werden vor dem Hintergrund von negativ verzerrten Leistungserwartungen und Bewertungen diskutiert (Glock et al., 2022; Tobisch & Dresel, 2017). Dementsprechend stellen sie Unsicherheitsfaktoren dar, da sie individuelle Charakteristika von Schüler*innen in den Hintergrund geraten lassen. Somit lässt sich die Untersuchung von Stereotypen sehr gut im Rahmen des Tagungsthema Tagungsthemas „Bildung zwischen Evidenz und Unsicherheit“ verorten.

Vor diesem Hintergrund befassen sich die Beiträge dieses interdisziplinären und multimethodischen Symposiums mit Stereotypen sowie mit Erwartungshaltungen und diagnostischen Prozessen, die durch Stereotype beeinflusst werden können. Die Beiträge richten unter Anwendung qualitativer und quantitativer Methoden den Blick auf unterschiedliche soziale Kategorien. Yendell et al. erheben in ihrer Mixed-Methods-Studie kulturelle Stereotype sowie individuelle Differenzkonstruktionen Lehramtsstudierender in Bezug auf unterschiedliche Armutsformen. Dabei fokussieren Sie auf die individuelle Perzeption überwiegend negativer kultureller Stereotype. Beißert et al. befassen sich in ihrer quantitativen Fragenbogenstudie mit herkunfts- und geschlechtsbezogenen Leistungserwartungen von Grundschullehrkräften im MINT-Bereich und finden, dass das Geschlecht, nicht aber die Herkunft eine relevante Kategorie für die Leistungseinschätzung darstellten. Kleen & Glock vereinbaren in ihrer Studie die Kategorien Geschlecht und Herkunft und finden, dass das Feedback von angehenden Lehrkräften für türkischstämmige Schülerinnen positiver ausfällt als für deutsche, während gleichzeitig türkischstämmige weniger mit Kompetenz assoziiert werden als deutsche. Schell et al. erfassen in Ihrer multimethodischen Studie Stereotype von angehenden Lehrer*innen gegenüber Schüler*innen mit Autismus-Spektrum-Störung, Down-Syndrom sowie Lernstörungen. Aus Interviews konnten insgesamt 35 relevante Stereotyp-Kategorien abgeleitet werden, die in einer anschließenden Umfrage quantifiziert und verifiziert wurden.

Die Beiträge werden von Anke Heyder eingeordnet und vor dem Hintergrund von Lehrer*innenbildung in allen Ausbildungsphasen diskutiert.

Beiträge des Symposiums

Kulturelle Stereotype und individuelle Differenzkonstruktionen angehender Lehrkräfte in Bezug auf unterschiedliche Armutsformen

Oscar Yendell[1], Carolina Claus[2], Meike Bonefeld[3], Karina Karst[1]
[1]Universität Mannheim, [2]Europa-Universität Flensburg, [3]Universität Konstanz

Tertiäre Herkunftseffekte werden neben primären und sekundären Herkunftseffekten herangezogen, um Mechanismen zur Entstehung von Bildungsungleichheit aufzudecken (Esser, 2016). Dabei wird das Handeln von (angehenden) Lehrkräften fokussiert sowie mögliche zugrundeliegende Stereotype, die als generalisierte Überzeugungen über soziale Gruppen verstanden werden (Eagly & Chaiken, 1993). Stereotype können in kulturell geteilte und individuelle Stereotype unterschieden werden (Eagly & Mladinic, 1989). Bisherige Studien zu Stereotypen von (angehenden) Lehrkräften auf die soziale Herkunft, fokussieren auf den sozioökonomischen Status (SES). In diesen Studien wird gezeigt, dass Schüler*innen eines niedrigen SES unter anderem im Hinblick auf Intelligenz, Ehrgeiz, und Arbeitsweise defizitärer eingeschätzt werden als Schüler*innen höherer SES (Dunkake & Schuchart, 2015; Tobisch & Dresel, 2020). Zugleich verweisen negative Stereotype auf negativ verzerrte Leistungserwartungen (Lorenz et al., 2016; Tobisch & Dresel, 2017). Perspektiven auf verschiedene Formen niedriger SES-Herkünfte bleiben aus. Dabei zeigen außerschulische Studien, dass Einkommensarme und Transferleistungsempfangende (bspw. „Hartz-4“) als Archetyp für negative Stereotype gelten (Asbrock, 2010; Cuddy et al., 2007) und Transferleistungsempfangende inkompetenter eingeschätzt werden als Einkommensarme (Suomi et al., 2020).

Fragestellung

In einer quantitativen Teilstudie wird untersucht, inwiefern sich Menge, Valenz und Inhalt wahrgenommener kultureller Stereotype von angehenden Lehrkräften bezogen auf Einkommensarme und Transferleistungsbeziehende unterscheiden. In einer qualitativen Teilstudie wird untersucht, ob und inwiefern angehende Lehrkräfte zwischen beiden Gruppen Differenzen auf Basis individueller Stereotype konstruieren (West & Fenstermaker, 1995). Übergreifend wird untersucht, wie kulturelle Stereotype individuell perzipiert werden.

Methode

Die Studie ist als konvergierende Mixed-Methods-Studie mit unterschiedlichen Samples angelegt (Creswell & Plano Clark, 2011). In der quantitativen Teilstudie nannten angehende Lehrkräfte (N=196) mittels offener Produktionsaufgabe kulturelle Stereotype jeweils für Einkommensarme und für Transferleistungsbeziehende. Diese wurden von zwei Ratern zu induktiv gebildeten Inhaltskategorien (Brennans-Prediger Kappa = 0.82) und negativen, neutralen sowie positiven Valenzen (Brennans-Prediger Kappa = 0.88) zugeordnet. Abschließend wurde eine mehrfaktorielle ANOVA mit Messwiederholung berechnet. Für die qualitative Teilstudie wurden problemzentrierte Interviews geführt (N=10) und nach der Grounded Theory ausgewertet (Corbin & Strauss, 1990; Witzel, 2000).

Ergebnisse

In der quantitativen Studie wird gezeigt, dass angehende Lehrkräfte mehr kulturelle Stereotype in Bezug auf Transferleistungsbeziehende als auf Einkommensarme nennen, $F(1, 195) = 11{,}27$, $p < .001$, $\eta 2 = .06$. Als Interaktionseffekt zeigt sich, dass mehr negative sowie weniger neutrale und positive kulturelle Stereotype gegenüber Transferleistungsbeziehenden als gegenüber Einkommensarmen geäußert werden, $F(1.27, 248.23) = 56.1$, $p < .001$, $\eta 2 = .22$. Für den Interaktionseffekt zwischen Gruppe, Valenz und Inhaltskategorie ($F(9.16, 1786.81) = 23.42$, $p < .001$, $\eta 2 = .11$) wurden Bonferroni-korrigierte paarweise Vergleiche gerechnet ($\alpha = .05$). Hier wird gezeigt, dass insbesondere bei individuellen Eigenschaften, signifikant mehr negative Stereotype gegenüber Transferleistungsempfangenden genannt werden als gegenüber Einkommensarmen (bspw. Leistungsbereitschaft und Verantwortungsbewusstsein).

In der qualitative Studie sehen alle Interviewten Transferleistungsempfangende öffentlich negativer konnotiert als Einkommensarme. Diese negative Konnotation bezieht sich auf (vermeintlich) individuelles Versagen (bspw. Faulheit). Einkommensarme sind in der Öffentlichkeit weniger sichtbar. Angehende Lehrkräfte, die dieser Konnotation folgen, differenzieren zwischen beiden Gruppen ursachenorientiert, indem sie Transferleistungsbeziehenden ein individuelles Versagen zuschreiben und bei Einkommensarmen auf strukturelles Versagen verweisen (bspw. politisches Versagen). Angehende Lehrkräfte mit persönlichem Kontakt zu Transferleistungsbeziehenden, folgen dieser ursachenorientierten Differenzierung nicht. Sie differenzieren auf beschreibender Ebene, indem sie Transferleistungsempfangende durch (vermeintliche) Arbeitslosigkeit und der Abhängigkeit von Transferleistungen desintegrierter als Einkommensarme beschreiben.

Übergreifend zeigt sich, dass negativere kulturelle Stereotype in Bezug auf die individuellen Eigenschaften von Transferleistungsempfänger*innen überwiegend die interviewten angehenden Lehrkräfte beeinflussen, die keinen Kontakt zu Transferleistungsbeziehenden haben. Sie können den negativen kulturellen Stereotypen durch fehlende persönliche Erfahrungen und fehlendes professionelles Wissen nichts entgegensetzen. Die Ergebnisse werden bezüglich weiterer Studienbedarfe diskutiert, die nicht mehr auf den niedrigen SES als Gesamtkonstrukt, sondern auf unterschiedliche niedrige SES-Herkünfte fokussieren und dabei mögliche Leistungserwartungseffekte, Bewertungseffekte sowie Attributionsprozesse untersuchen.

Stereotype Überzeugungen? Geschlechts- und herkunftsbezogene Erwartungshaltungen von Lehrkräften im MINT-Bereich

Hanna Beißert[1], Luke McGuire[2], Kelly Lynn Mulvey[3]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, [2]University of Exeter, [3]North Carolina State University

Ausgangspunkt der Studie ist, dass im mathematischen, technischen und naturwissenschaftlichen Bereich Mädchen (Quaiser-Pohl & Endepohls-Ulpe, 2010) und Kinder mit Migrationshintergrund (Estrada et al., 2011) häufig unterrepräsentiert sind. In diesem Zusammenhang wird immer wieder gezeigt, dass Mädchen weniger Freude und geringeres Interesse am MINT-Bereich zeigen, und dass Mädchen sowie Kinder mit Migrationshintergrund ungünstigere leistungsbezogene Kognitionen und Emotionen (z.B. negativere Fähigkeitsselbstkonzepte, höhere Prüfungsangst) aufweisen, und sich

seltener für Kurse und Karrierewege im MINT-Bereich entscheiden (z.B. Lazarides & Ittel, 2017; Martiny et al., 2012; Museus et al., 2011; Quaiser-Pohl, 2012; Estrada et al., 2011).

Ein Ansatzpunkt, um solche geschlechts- und herkunftsbezogenen Tendenzen zu erklären, sind Erwartungshaltungen durch stereotype Überzeugungen. Die stereotypkonforme Wahrnehmung des MINT-Bereichs als männliche Domäne sowie niedrigere Erwartungen bezüglich Fähigkeiten, Leistungen und Interessen von Mädchen oder Kindern mit Migrationshintergrund im MINT-Bereich können die Lehrkräfte in ihrer Wahrnehmung der Schüler*innen beeinflussen und sich auf ihr konkretes Unterrichtsverhalten sowie auf Förder- und Auswahlentscheidungen auswirken (Glock & Kleen, 2019; Jussim et al., 1996).

Die vorliegende Studie untersucht, inwiefern sich stereotype Sichtweisen in den Erwartungshaltungen von Grundschullehrkräften niederschlagen und ob dabei Domänenunterschiede innerhalb des MINT-Bereichs bestehen. Zudem werden Zusammenhänge von stereotypen Erwartungen mit den impliziten Fähigkeitstheorien der Lehrkräfte untersucht.

Methode

An der Studie nahmen 43 Klassenlehrkräfte von dritten oder vierten Klassen aus Grundschulen in Baden-Württemberg teil (37 weiblich, 6 männlich, M=45.07 Jahre, SD=11.30, 27-65 Jahre). Anhand von Fragebögen wurden die Erwartungshaltungen der Lehrkräfte bezüglich der Leistungen von Mädchen, Jungen, deutschen Kindern und türkischen Kindern in verschiedenen Domänen des MINT-Bereichs (Mathematik, Naturwissenschaften, Technik, Informatik) erfasst. Zur Angabe dieser Einschätzungen standen den Lehrkräften Zahlenstrahle (0-100) zur Verfügung. Darüber hinaus wurden die impliziten Fähigkeitstheorien der Lehrkräfte anhand der Skala von Dresel et al. (2015) erfasst.

Ergebnisse

Sowohl um die Leistungseinschätzungen bezüglich Jungen im Vergleich zu Mädchen zu betrachten als auch für den Vergleich der Leistungseinschätzungen bezüglich deutscher versus türkischer Kinder, wurde jeweils eine Messwiederholungs-ANOVA gerechnet mit den verschiedenen MINT-Domänen sowie den verschiedenen Kindergruppen als Messwiederholungsfaktoren.

Wie erwartet zeigte sich ein Haupteffekt für das Geschlecht der Kinder ($F(1, 37)=14.61$, $p<.001$, $\eta p^2=.28$) sowie eine Interaktion aus dem Geschlecht und der spezifischen Domäne des MINT-Bereichs $F(1.80, 66.48)=7.30$, $p=.002$, $\eta p^2=.17$. Konkret wurden die Fähigkeiten von Jungen in den Bereichen Mathematik, Technik und Informatik besser eingeschätzt als die der Mädchen, $ps<.01$. Im Bereich Naturwissenschaften wurde der Geschlechtsunterschied nicht signifikant, $p=.052$. Am größten waren die Unterschiede in den Bereichen Technik und Informatik – zwei Bereiche, die traditionell stark stereotypisiert sind.

Im Hinblick auf die Herkunft der Schüler*innen (deutsche Kinder vs. türkische Kinder) zeigten sich keine Unterschiede in den Leistungserwartungen der Lehrkräfte. Die kulturelle Herkunft stellte also in der vorliegenden Stichprobe keine relevante soziale Kategorie für die Leistungseinschätzung dar. Da frühere Forschungsarbeiten aber Zusammenhänge von Leistungsbeurteilungen mit der kulturellen

Herkunft gefunden haben (z.B. Bonefeld et al., 2017), sollte hinterfragt werden, ob die Herkunft für die Leistungseinschätzung von Lehrkräften im MINT-Bereich wirklich nicht relevant war oder ob hier möglicherweise Effekte sozialer Erwünschtheit oder stärkerer gesellschaftlicher Sensibilisierung vorliegen.

Im Hinblick auf die impliziten Fähigkeitstheorien der Lehrkräfte zeigte sich, dass ein stärker ausgeprägtes „growth mindset“ mit geringeren Unterschieden in den erwarteten Leistungen von Jungen vs. Mädchen einherging.

Diskussion

Insgesamt lässt sich festhalten, dass die Einschätzungen der Lehrkräfte bezüglich der Fähigkeiten von Mädchen und Jungen im MINT-Bereich entsprechend gängigen Geschlechtsstereotypen ausfallen, wobei dies in geringerem Maße der Fall ist, wenn die Lehrkräfte ein stärker ausgeprägtes Growth-Mindset aufweisen. Dementsprechend könnten Interventionsmaßnahmen an den impliziten Fähigkeitstheorien der Lehrkräfte ansetzen. Zudem sollte auf die Sensibilisierung von Lehrkräften abgezielt werden, damit diese sich ihrer eigenen stereotypen Überzeugungen bewusstwerden und den Auswirkungen auf Unterrichtshandeln oder Förderentscheidungen aktiv entgegenwirken können, um allen Kindern unabhängig ihres Geschlechts eine Förderung zu ermöglichen, die ihren Fähigkeiten entspricht.

Hausfrau oder Karrierefrau? Unterfordert oder überfordert? - Stereotype Assoziationen von angehenden Lehrkräften gegenüber Schülerinnen mit türkischem Migrationshintergrund

Hannah Kleen[1], Sabine Glock[2]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, [2]Bergische Universität Wuppertal

In dieser Studie werden stereotype Assoziationen gegenüber türkischstämmigen Schülerinnen untersucht. Dass türkischstämmige Schüler*innen in Deutschland benachteiligt sind, ist hinlänglich bekannt. Die Rolle des Geschlechts, welches bei deutschen Schüler*innen durchaus eine Relevanz im Bildungsweg hat, wurde innerhalb der Gruppe der türkischstämmigen Schüler*innen bisher jedoch nur zum Teil betrachtet, wobei sich hier ein differenzielles Bild zeigt. So werden türkischstämmige Schülerinnen in weiterführenden Schulen implizit negativer wahrgenommen als türkischstämmige Schüler (Glock & Klapproth, 2017). Dies wird unter anderem damit begründet, dass Lehrkräfte davon ausgehen könnten, türkischstämmige Schülerinnen unterlägen bestimmten Geschlechterrollen und seien weniger an einem hohen Schulabschluss interessiert als türkischstämmige Schüler. Weiterhin sind türkischstämmige Schülerinnen je nach Leistungsdimension benachteiligt, was sowohl auf die Ethnie als auch auf das Geschlecht zurückgeführt wird (Bonefeld et al., 2021). Es gibt jedoch erste Hinweise, dass geschlechtsstereotype Zuschreibungen eher die Schüler*innen betrifft, die der ethnischen Majorität angehören als Schüler*innen einer ethnischen Minderheit (Bonefeld et al., 2021).

Stereotype beschreiben generalisiertes Wissen und Überzeugungen über bestimmte soziale Gruppen, die Verhalten beeinflussen können (Devine, 1989). Stereotype Inhalte werden dem Stereotype Content Model (Fiske et al., 2002) zufolge in Wärme- und Kompetenzstereotype unterteilt. Beide Inhaltsbereiche spielen in der zwischenmenschliche Interaktion eine wichtige Rolle (Fiske et al., 2002) und es lassen sich dort verschiedene soziale Gruppen einordnen (Asbrock, 2010). Obwohl türkischstämmige Schülerinnen zum Teil negativer wahrgenommen und schlechter beurteilt werden

als türkischstämmige Schüler, ist über die stereotypen Assoziationen von Lehrkräften über speziell türkischstämmige Schülerinnen bisher wenig bekannt. So wurden in einer Studie beispielsweise Stereotype gegenüber türkischstämmigen Personen erhoben, aber nicht weiter nach Geschlecht unterschieden (Bonefeld & Karst, 2020).

Das Ziel diese Studie war es daher, sowohl stereotype Inhalte (Fiske et al., 2002) zu erheben, als auch weitere stereotype Assoziationen von angehenden Lehrkräften über türkischstämmige Schülerinnen. Im Stereotype Content Model wurde dabei zusätzlich—mit Bezug auf implizit negative Einstellungen von Sekundarschullehrkräften gegenüber türkischstämmigen Schülerinnen—eine Unterscheidung zwischen Haus- und Karrierefrau hergestellt, um einen möglichen Einfluss von Geschlechterrollen mit abzubilden.

Insgesamt haben 79 Lehramtsstudierende (61 weiblich, MAlter=25,03 Jahre, SDAlter=2,34) an der Studie teilgenommen. Die Erhebung der Wärme- und Kompetenzstereotype hatte ein 2x2 between-subjects Design mit den Faktoren Ethnie (türkischstämmig vs. deutsch) und Geschlechterrolle (Hausfrau vs. Karrierefrau). Für die Erfassung der stereotypen Assoziationen von angehenden Lehrkräften wurde ein einfaktorielles (deutsch vs. türkischstämmig) between-subjects Design verwendet. Die Lehramtsstudierenden erhielten eine leistungsmäßig ambige Schülerinnenbeschreibung und schrieben daraufhin ein Feedback in einem offenen Format.

Mittels MANOVA mit den Faktoren Ethnie (türkischstämmig vs. deutsch) und Geschlechterrolle (Hausfrau vs. Karrierefrau) wurden die Stereotypdimensionen Wärme und Kompetenz auf Signifikanz geprüft. Die beiden signifikanten Haupteffekte wurden durch eine signifikante Interaktion, $F(1,72)=9,28$; $p<.001$, qualifiziert. Post-hoc Tests zeigten, dass innerhalb der Gruppe der Karrierefrauen, türkischstämmige als signifikant wärmer als deutsche eingeschätzt wurden, während es bei der Kompetenz der umgekehrte Fall war. Innerhalb der Gruppe der Hausfrauen gab es keine signifikanten Unterschiede zwischen deutschen und türkischstämmigen Personen. Interessanterweise zeigte sich innerhalb der deutschen Personen, dass Hausfrauen als signifikant wärmer, jedoch weniger kompetent eingeschätzt wurden als Karrierefrauen, während sich keine signifikanten Unterschiede bei den türkischstämmigen Personen zeigten.

Die offenen Antworten wurden mittels automatisierter Inhaltsanalyse ausgewertet. Dabei wurde einerseits anhand der Textmetriken ausgewertet, welche Wörter jeweils (d.h. türkischstämmige vs. deutsche Schülerin) am häufigsten verwendet wurden. Hier zeigt sich, dass beispielsweise der Name der Schülerin häufiger in der deutschen als in der türkischstämmigen Bedingung genannt wurde. Ebenso war es bei den Wörtern „Unterrichtsstörung“, „Eltern“ und „unterfordert“. Beim Feedback der türkischstämmigen Schülerin wurden zum Beispiel häufiger die Wörter „üben“, „fördern“ und „unterstützen“ genannt. Weiterhin wurde eine automatisierte Sentiment Analyse durchgeführt. Die Ergebnisse verdeutlichen, dass das Feedback der deutschen Schülerin mehr negativ konnotierte Wörter enthielt als das Feedback der türkischstämmigen Schülerin.

Die vollständigen Ergebnisse werden im Rahmen des Vortrags aufgezeigt und diskutiert.

Judging a book by its cover?! Stereotype von Lehrkräften gegenüber Kindern mit Förderbedarf

Charlotte Schell[1], Charlotte Dignath[2], Nathalie John[1], Mareike Kunter[1]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, [2]Technische Universität Dortmund

Theoretischer Hintergrund

Lehrkräfte müssen regelmäßig den heterogenen Lernfortschritt ihrer Schüler*innen diagnostizieren, um pädagogische und didaktische Entscheidungen abzuleiten (Helmke, 2012). Die Forschung zeigt, dass Lehrkräfte hierbei eher zu heuristischen Urteilsprozessen neigen, die zwar schnell und effizient, aber auch weniger reflektiert und damit anfällig für Verzerrungen sind (Krolak-Schwerdt et al., 2018). Hierbei können Stereotype über die Schüler*innen eine Rolle spielen (Bonefeld & Karst, 2020). Stereotype sind Überzeugungen über Eigenschaften und Verhaltensweisen von Mitgliedern bestimmter sozialer Gruppen, die die Informationsverarbeitung beeinflussen können (Hilton & von Hippel, 1996). Im inklusiven Kontext sind sie besonders relevant, denn Lehrkräfte bewerten die Leistungen von Schüler*innen mit Förderbedarf schlechter im Vergleich zu Schüler*innen ohne Förderbedarf (Pit-ten Cate & Glock, 2019; Pit-ten Cate & Krischler, 2020). Typische Eigenschaften, die Menschen mit Förderbedarf zugeschrieben werden, sind Inkompetenz (Rohmer & Louvet, 2012), Abhängigkeit (Popovich et al., 2003), aber auch Wärme (Fiske et al., 2012).

Zusätzlich zu diesen allgemeinen Tendenzen gibt es spezifische Stereotypen gegenüber Gruppen von Schüler*innen mit bestimmten Individuellen Förderbedarfen (IF; Glock et al., 2020). Zur Beschreibung von Stereotypen wird das Stereotype Content Models von Fiske et al. (2002) verwendet, welches Stereotype als Kombination der zwei Dimensionen Wärme und Kompetenz beschreibt. So werden Kinder mit Down-Syndrom (DS) als warm, aber nicht sehr kompetent wahrgenommen (Fiske, 2012). Menschen im Autismus-Spektrum (AS) hingegen werden als weniger warmherzig wahrgenommen (Fiske, 2012) und mit Savant-Fähigkeiten assoziiert (Bennett et al., 2018). Schüler*innen, die von Lernstörungen (LS) aufweisen, werden als weniger kompetent und leistungsfähig wahrgenommen, als sie tatsächlich sind (Pit-ten Cate & Krischler, 2020).

Fragestellung

Aufbauend auf den bisherigen Befunden prüfen wir in zwei Teilstudien folgende Hypothesen:

1. Lehramtsstudierende assoziieren verschiedene Stereotype mit den Kindern mit AS, DS und LS.

2. Diese lassen sich in das Stereotype Content Model einordnen.

Methode

In Studie A wurden N=13 Lehramtsstudierende anhand eines strukturierten Interviews mit größtenteils offenen Fragen zu ihren Stereotypen über die drei Gruppen von Schüler*innen mit IF befragt. In Studie B wurden in Kombination mit der deutschen Übersetzung (Krischler et al., 2018) des bereits etablierten Stereotype Content Fragebogen (Fiske et al. 2002) die relevantesten der in Studie A genannten expliziten Stereotypen mittels eines Onlinefragebogens an einer größeren heterogeneren Stichprobe von N=213 Lehramtsstudierenden aus ganz Deutschland quantifiziert und verifiziert.

Ergebnisse

Die Interviews aus Studie A wurden aufgezeichnet, transkribiert und anhand der qualitativen Inhaltsanalyse (Kuckartz & Rädiker, 2022) kodiert und ausgewertet. Als besonders relevant wurden 15 Kategorien zu ASS, 12 zu DS und 8 zu LS identifiziert. Hierzu gehörten zum Beispiel, bezogen auf

ASS „inselbegabt“, „introvertiert“ und „impulsiv“. In Bezug auf DS waren „offen“, „leistungsschwach“ und „warmherzig“ wichtig. Auch zu LS wurde „leistungsschwach“ genannt, sowie „nicht intelligent“ und „nicht frustrationstolerant“.

Studie B: Um die Dimensionen Wärme und Kompetenz zu bilden, wurden die Mittelwerte über die jeweiligen vier Wärme- und Kompetenzadjektive analog zu Fiske et al. (2002) gebildet. Mittels t-Tests wurden diese Dimensionen erst einmal mit dem Mittelwert der Skala (3.5) verglichen. Hier unterschieden sich alle Einschätzungen signifikant vom Mittelwert mit Ausnahme der Dimension Kompetenz bei ASS.

Eine rmMANOVA mit dem within-subject Faktor Förderbedarf und den Dimensionen Wärme und Kompetenz als AV wurde signifikant, $F(4,818)=250.17$, $p<.001$, $np2=.550$. Paarweise t-Tests zeigten, dass Schüler*innen im AS als weniger warm eingeschätzt wurden als Schüler*innen mit DS oder mit LS. Darüber hinaus wurden Schüler*innen mit DS als wärmer eingestuft als Schüler*innen mit LS. Bei der Kompetenz war es umgekehrt. Schüler*innen mit DS wurden zudem als weniger kompetent eingestuft als Schüler*innen mit LS. Schließlich wurden Schüler*innen im AS eher kompetenter als wärmer eingestuft, wohingegen Schüler*innen mit DS und mit LS eher wärmer als kompetent eingestuft wurden

Weitere Ergebnisse zu den abgefragten Adjektiven werden momentan analysiert und sollen auf der GEBF-Konferenz präsentiert und diskutiert werden.

G-15: Monitoring nachgeordneter Bildungsbereiche in Zeiten gesellschaftlicher Unsicherheit. Ansätze und Instrumente zur Erfassung der politischen und kulturellen Bildung sowie zur Erfassung der Bildung für nachhaltige Entwicklung

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* T03 R03 D75

Monitoring nachgeordneter Bildungsbereiche in Zeiten gesellschaftlicher Unsicherheit. Ansätze und Instrumente zur Erfassung der politischen und kulturellen Bildung sowie zur Erfassung der Bildung für nachhaltige Entwicklung

Chair(s): **Gerhard de Haan** (Institut Futur)
*Diskutant*in(nen):* **Nina Jude** (Universität Heidelberg)

Bildungsberichterstattung ist der Versuch, ein Bildungssystem in seinen Institutionalisierungen, Prozessen und Ergebnissen im Zeitverlauf so abzubilden, dass eine verlässliche Grundlage für Entscheidungen zu Veränderung oder Fortführung des Bestehenden gegeben wird. Welche Kompetenzbereiche und Aspekte eines Bildungssystems dabei genau in den Blick kommen, ist abhängig von der verfügbaren Datengrundlage, von wissenschaftlichen Diskursen (Klieme et al., 2006; Maaz & Kühne, 2018; Grünkorn et al. 2019) und politischen Entscheidungen. Auf der politischen Ebene legt für Deutschland die Kultusministerkonferenz in ihrer Gesamtstrategie zum Bildungsmonitoring (KMK 2006; 2015) fest, welche Kompetenzbereiche des Bildungssystems zum Gegenstand des Bildungsmonitorings auf der föderalen Ebene werden sollen; es sind dies Lesen, Mathematik, Naturwissenschaften und die erste Fremdsprache (KMK, 2015, 7-9).

Diese exklusive Fokussierung ist schon vielfach kritisiert worden (Schrader & Herbrechter, 2015; Abs, 2009), weil somit die Aufmerksamkeit in Lehrerbildung, Schulentwicklung sowie hinsichtlich der Unterstützungsagenturen einseitig gesteuert und in Kauf genommen wird, dass andere Kompetenzbereiche in den Hintergrund treten. Dies rächt sich besonders dann, wenn in Phasen gesellschaftlicher Unsicherheit und Transformation, bislang nachgeordnete Kompetenzbereiche an Bedeutung gewinnen. Die gegenwärtige gesellschaftliche Unsicherheit ist durch eine ökologische Krise gekennzeichnet, die sich mit einer Krise des sozialen Zusammenhalts und der Unterstützung des politischen Systems verbindet (Abs, 2021). Um auf die anstehenden Transformationsherausforderungen im Bildungssystem reagieren zu können, ist deshalb ein Monitoring von Kompetenzbereichen erforderlich, die diese Krisen adressieren.

Akteure der kulturellen Bildung, der politischen Bildung und der Bildung für nachhaltige Entwicklung haben erkannt, dass belastbare Daten zur systematischen und kontinuierlichen Erfassung der Voraussetzungen und Leistungsfähigkeit in diesen Kompetenzbereichen fehlen und ein Bildungsmonitoring entsprechend weiterentwickelt werden muss (z.B. Becker 2011; Deutscher Bundestag 2021: 21; 142; Rat für kulturelle Bildung 2022).

Vor diesem Hintergrund werden im Rahmen des Symposiums indikatorenbasierte Monitoringansätze und -strategien für die drei genannten Bildungsbereiche vorgestellt und diskutiert. Der erste Artikel fokussiert dabei auf den Beitrag von Kultureinrichtungen und anderen non- und informellen Lerngelegenheiten zur kulturellen Bildung, gibt Einblicke in die zum Teil prekäre Datenlage und stellt auf dieser Basis ausgewählte Indikatoren vor. Der zweite Artikel stellt Ergebnisse einer multi-methodischen Messung der Implementation von BNE im deutschen Bildungssystem in den Mittelpunkt und führt Erkenntnisse zu treibenden und hemmenden Kräften aus. Der dritte und vierte

Beitrag nehmen anschließend das Feld der politischen Bildung aus verschiedenen Perspektiven in den Blick: Ersterer untersucht die Implementierung politischer Bildung im Rahmen der hochschulischen Lehrkräfteausbildung, letzterer Möglichkeiten und Grenzen eines datenbasierten Monitorings der außerschulischen politischen Bildung am Beispiel entwickelter Indikatoren für den Akteur Volkshochschulen.

Beiträge des Symposiums

Entwicklung und Ergebnisse eines indikatorengestützten Monitorings für das Feld der kulturellen Bildung in Deutschland

Margot Belet, Jannis Burkhard, Katharina Nuss, Juliane Baumann, Stefan Kühne, Kai Maaz
Leibniz-Institut für Bildungsforschung und Bildungsinformation

Theoretischer Hintergrund

Die nationale Berichterstattung fokussiert auf die institutionellen Angebote in den jeweiligen Bildungsstufen, der Bildungserwerb in anderen Settings wird nach neun nationalen Bildungsberichten vergleichsweise weniger stark erfasst. Dies betrifft auch die kulturelle Bildung, trotz zunehmender politischer Aufmerksamkeit in den letzten Jahren (vgl. z. B. KMK 2013). Ein Grund für die mangelnde Datenlage ist die Breite des Feldes mit seinen dezentralen Stakeholder*innen: Neben formalen Bildungseinrichtungen leisten Kultureinrichtungen und viele weitere informelle und non-formale Lernumwelten einen wichtigen Beitrag. 2012 wurde die „kulturelle/musisch-ästhetische Bildung im Lebenslauf“ als Schwerpunkthema im nationalen Bildungsbericht bearbeitet. Dies lieferte erste Beschreibungen und auch wenige sozialstrukturelle Differenzierungen. Das Thema wurde jedoch nachfolgend weder in der Bildungsberichterstattung noch in anderen Monitoringinstrumenten systematisch aufgenommen.

Der Rat für Kulturelle Bildung (2017, 2019) hat sich seit 2013 zur Lage und Qualität der kulturellen Bildung in Deutschland geäußert. Auch diese zentralen Untersuchungen beschränkten sich aber auf spezifische Feldausschnitte zu einem bestimmten Zeitpunkt. Die Notwendigkeit eines dauerhaften Monitorings wurde offensichtlich, um die beschriebenen Bedarfslagen längsschnittlich verfolgen und darauf aufbauend Steuerungshandeln vorbereiten zu können (s. auch Maaz & Kühne 2018).

Methode

Die Monitoringentwicklung ist das Ziel des vom DIPF durchgeführten und von Stiftung Mercator geförderten Projekts „InKuBi – Indikatoren für kulturelle Bildung“. InKuBi führt keine neuen Erhebungen durch, sondern sichtet vorhandene Datenquellen und analysiert diese systematisch. Dafür wurde kulturelle Bildung als konkreter Beobachtungsgegenstand fokussiert. Die breite Begriffsdefinition des Projekts sieht kulturelle Bildung als Teil lebenslanger Allgemeinbildung und geht von einem weiten Kulturverständnis aus, das von den Künsten im engen Sinne bis zu weiteren (sozio-)kulturellen, ästhetisch-expressiven Praktiken reicht.

Diesem Begriffsverständnis wurde sich aus drei Perspektiven angenähert: Die Angebotsperspektive beschreibt die Gelegenheitsstrukturen kultureller Bildung, die Prozessperspektive rückt die Nutzung kultureller Bildung in den Mittelpunkt, und die Ergebnisperspektive beschreibt die Wirkungen kultureller Bildung auf individueller sowie gesellschaftlicher Ebene.

Innerhalb dieser drei Perspektiven wurde zunächst theoretisch hergeleitet, welche Themen für das Feld der kulturellen Bildung so entscheidend sind, dass sie in einem Monitoring abgebildet werden sollten. Dann wurde die konkrete Operationalisierung in Form von Indikatoren und Kennziffern vorgenommen. Angelehnt an die nationale Bildungsberichterstattung liegt ein weiter Indikatorbegriff zugrunde, sodass sich ein Indikator aus mehreren Kennziffern zusammensetzt, die argumentativ zueinander ins Verhältnis gesetzt werden. Da die Fülle an Indikatoren im Rahmen des Symposiums nicht dargestellt werden kann, wird exemplarisch die Erarbeitung eines Indikators aus dem Bereich der Gelegenheitsstrukturen erläutert.

Ergebnisse

Als Beispiel wird im Folgenden der Indikator „Strukturelle Rahmenbedingungen für kulturelle Bildung an allgemeinbildenden Schulen der Sekundarstufe II“ vorgestellt. Dieser betrachtet die personelle Ausstattung sowie die Profilbildung der Schulen. Alle Analysen dafür wurden auf Basis von Daten aus der Schulleitungsbefragung der Startkohorte 3 aus dem Nationalen Bildungspanel (NEPS) vorgenommen, die im Jahr 2015 erhoben.

Im Durchschnitt stehen einer Schule der Sekundarstufe II in Deutschland 4,58 Lehrkräfte (VZÄ) pro Schulfach zur Verfügung. Die künstlerischen Fächer liegen mit 3,9 VZÄ (Kunst) bzw. 3,2 VZÄ (Musik) pro Schule im unterdurchschnittlichen Bereich. Im Vergleich der Schularten ergibt sich das Bild, dass insbesondere die Hauptschulen für das Fach Musik personell nicht gut ausgestattet sind: An 11,7 % der Hauptschulen arbeitet keine Lehrkraft im Fach Musik. Im Kontrast dazu steht allen Gymnasien sowie allen integrierten Gesamtschulen mindestens eine Musiklehrkraft zur Verfügung. Damit einher geht der Befund, dass Gymnasien und integrierte Gesamtschulen öfter als die anderen Schularten musische Profile bilden. Insgesamt liegt der Anteil der Schulen mit musischem Schwerpunkt bei 13,9 % aller Schulen.

Anhand der Ergebnisse des Beispielindikators diskutiert der Beitrag auch die eingeschränkte Datenverfügbarkeit, die theoretische Eingrenzung des Untersuchungsgegenstands, die Steuerungsrelevanz der Ergebnisse sowie die Zukunftsperspektive des Monitorings. Der Beitrag greift also neben der Entwicklung des Monitors und dessen inhaltlicher Ausgestaltung unter Einbezug der anderen Beiträge des Symposiums auch übergreifende Herausforderungen auf.

Monitoring von Bildung für nachhaltige Entwicklung in Deutschland

Antje Brock, Gerhard de Haan, Julius Grund, Jorrit Holst, Mandy Singer-Brodowski
Institut Futur

Theoretischer Hintergrund

Monitoringverfahren im Bildungssystem sind seit den PISA-Studien mit dem Ziel verknüpft, auf Basis einer systematischen Beobachtung die Evidenz zu liefern, auf deren Grundlage Qualität und Leistungsfähigkeit des Bildungssystems gesichert und erhöht werden soll (Grünkorn et al. 2019).

Für Deutschland wird seit 2015 die Qualität und Quantität von nachhaltigkeitsrelevanter Bildung im Bildungssystem umfänglich und systematisch in einem vom BMBF geförderten Forschungsprojekt erfasst. Es handelt sich um ein unabhängiges Monitoring an der Schnittstelle von Wissenschaft und Politik und orientiert sich an internationalen Empfehlungen zur kontextualisierten und multi-

methodischen Erfassung von BNE in Monitoring- und Evaluationsverfahren (Stepanek Lockhardt 2018). Durchgeführt wird es in der Arbeitsstelle des wissenschaftlichen Beraters der Nationalen Plattform zur Umsetzung des UNESCO-Programms ESD for 2030 in Deutschland (FU Berlin, Institut Futur).

Methode

International wird aktuell zur Indikatorisierung von BNE häufig die Indikatorik der Sustainable Development Goals (SDGs) verwendet, hier im Besonderen der Indikator SDG 4.7 im Kontext des Ziels 4 “Hochwertige Bildung” (OECD 2022). Im Kontrast zu solchen Indikatoren, die auf der Basis von Selbstauskünften verantwortlicher Institutionen bewertet werden, liegen dem BNE-Monitoring in Deutschland verschiedene methodische Studien zugrunde: Dokumentenanalysen, qualitative Studien und quantitative Studien.

Jeweils im zeitlichen Abstand von höchstens drei Jahren wird eine lexikalische Dokumentenanalyse für die vier formalen Bildungsbereiche Frühkindliche Bildung, Schule, Berufliche Bildung und Hochschule durchgeführt, um die Implementierung von BNE auf der Input-Ebene des Bildungssystems längsschnittlich zu erfassen (Singer-Brodowski et al. 2019; Holst et al. 2020). Bei mittlerweile über 11.000 analysierten Dokumenten zeichnet sich ein erkennbarer, jedoch langsamer Trend der Aufnahme von BNE und verwandten Bildungskonzepten in zentralen Dokumenten ab (z.B. in Bildungspläne, Curricula, Ausbildungsordnungen, Studiengangsordnungen). Zugleich stellen gerade die Dokumente zur Aus- und Weiterbildung von pädagogischem Personal, also frühpädagogischen Fachkräften, Lehrkräften für allgemeinbildende und berufliche Schulen, sowie ausgewählte Dokumente hochschuldidaktischer Weiterbildungen über die Zeitreihe hinweg ein wesentliches Desiderat dar (Singer-Brodowski & Holst 2022; Brock & Holst 2022; Holst 2022; Holst & Singer-Brodowski 2022).

Auf der Outcome-Ebene wurden durch zwei großangelegte Online-Access-Panel-Studien in 2018 und 2020 (jeweils n > 3.000) verschiedene Bereiche der BNE-Implementierung quer- und längsschnittlich erhoben. Hierzu zählen u. A. nachhaltigkeitsrelevantes Wissen, Einstellungen, Emotionen, BNE-Implementierung sowie selbst berichtetes Verhalten von jungen Menschen (14-24 Jahre) und Lehrkräften (Brock & Grund 2018; Grund & Brock, 2018; Grund & Brock in Vorbereitung). Die umfangreichen Datenreihen verweisen auf Ebene der jungen Menschen darauf, dass sich die Berührungspunkte zu Nachhaltigkeit (z.B. Medien, Freundeskreis, Freizeit) binnen der letzten vier Jahre in allen Lebensbereichen deutlich erhöht haben, außer in der formalen Bildung (ebd). Gleichzeitig gaben die jungen Menschen wie auch die Lehrkräfte an, dass sie sich eine wesentlich intensivere Auseinandersetzung mit Nachhaltigkeit in den Bildungseinrichtungen wünschen – idealerweise in mehr als 30 % ihrer Unterrichts-/Seminarzeit. Äquivalent zur Dokumentenanalyse geht auch aus den Lehrkräftebefragungen hervor, dass insbesondere die Befähigung zu einer guten BNE-Umsetzung in der Ausbildung von Lehrenden ein zentrales Desiderat ist (ebd.). Wird BNE umgesetzt, ist dies zumeist stark fächerspezifisch der Fall (vor allem in thematisch affinen “Trägerfächern” wie Geographie, Biologie oder Sozialkunde) und abhängig von einzelnen interessierten Lehrenden. Die Studien verdeutlichen zudem ein insgesamt hohes Problembewusstsein unter Lernenden und

Lehrenden bezogen auf verschiedene Nachhaltigkeitsbereiche. Die Einschätzung von Zukunftsszenarien zeigt einen verbreiteten Zukunftspessimismus in beiden Gruppen auf, etwa in Bezug auf die effektive Bearbeitung des Klimawandels (Grund & Brock 2019).

Ergebnisse

Die Ergebnisse der verschiedenen Teilstudien des Monitorings, die im Rahmen einer Methodentriangulation regelmäßig systematisch aufeinander bezogen werden, geben Einblicke in die tieferen Treiber und Hemmnisse der Implementation von BNE im deutschen Bildungssystem. Sie leisten darüber hinaus zentrale Beiträge für eine weitere Aufbereitung BNE-bezogener Erhebungsverfahren für längsschnittlich angelegte Monitorings und die Bildungsberichterstattung.

Indikatorenentwicklung für ein Monitoring politischer Bildung an der Hochschule

Lucy Huschle, **Hermann Josef Abs**
Universität Duisburg-Essen

Theoretischer Hintergrund

Bestehende Berichtssysteme zur politischen Bildung sind international initiiert und fokussieren den schulischen Bildungsbereich (vgl. European Commission 2018; Schulz et al. 2018). Hingegen fehlen nationale Berichte dazu, welche Aktivitäten politischer Bildung an der Hochschule stattfinden. Dabei wurde die öffentliche Debatte um das Verhältnis und die Bedeutung des Bildungsortes Hochschule für die Konstitution der Demokratie und die Persistenz einer freien Gesellschaft in Deutschland schon im Vorfeld der Studierendenproteste in den 60iger-Jahren geführt (vgl. z. B. Nitsch et al. 1965). Auch dokumentieren vielfältige Forschungen Einflüsse der allgemeinen Hochschulbildung auf demokratische Einstellungen, Fähigkeiten und Verhaltensweisen von Studierenden. Im Einzelnen werden Einflüsse auf die politische Partizipation, das gesellschaftliche Engagement oder die politischen Orientierungen analysiert (Yang & Hoskins 2020; Celio et al. 2011; Havey & Schalewski 2022). Besonderes Gewicht kommt der Hochschule dann zu, wenn sie formale Programme zur politischen Bildung verantwortet (Europarat 2020; Tierney 2021; Weschenfeld 2014). Auf systematische Weise geschieht dies im Bereich der Lehrkräftebildung für den politisch bildenden Unterricht an Schulen. Aber nicht nur die Bildungsprogramme für künftige Politiklehrkräfte sind hier relevant, sondern alle Lehrkräfte kommen in den Blick, wenn diese als „Agenten der Demokratie“ (vgl. BMBF 2022) betrachtet werden, deren Kernaufgabe darin besteht, ihre erworbenen demokratischen Kompetenzen in Prozessen der politischen Bildung weiterzugeben. Lehramtsstudiengänge – ihre strukturellen Voraussetzungen, curricularen Inhalte und ihre Rezeption durch Studierende – bilden damit im Feld der hochschulischen politischen Bildung einen Untersuchungsgegenstand mit besonderer Relevanz.

Vor dem skizzierten Hintergrund lautet die leitende Forschungsfrage dieses Vortrags: Wie lässt sich die politische Bildung an Hochschulen im Rahmen der Lehramtsstudiengänge systematisch, repräsentativ und wiederkehrend erfassen?

Methode

Zur Bearbeitung der Fragestellung wurde zunächst eine Matrix zur Erfassung von politischer Bildung an Hochschulen in Deutschland entwickelt. Davon ausgehend wurde nach geeigneten

Datenbeständen in der amtlichen Statistik, öffentlich zugänglichen Dokumenten, und repräsentativen Befragungen gesucht. Der Vortrag zeigt beispielhafte Zugänge und Analysen, die zu verschiedenen Aspekten des Lehramtsstudiums im Zusammenhang mit politischer Bildung entwickelt wurden.

Genauer werden drei konzipierte Indikatoren näher betrachtet: Auf der Input-Ebene wird ein Indikator vorgestellt, der die Forschungsbasis der politikdidaktischen Forschung an der Hochschullehre misst. Auf der Prozessebene wird ein Indikator ausgeführt, der Anteile politischer Bildung im bildungswissenschaftlichen Anteil des allgemeinen Lehramtsstudiums erfasst. Auf der Ergebnisebene wird ein Indikator vorgestellt, der die Studierenden- und Absolvierendenzahlen im Lehramtsstudium mit dem aktuellen Bedarf an neuen Politiklehrkräften in Beziehung setzt, um so mögliche Defizite aufzudecken. Entlang der Präsentation der drei beispielhaften analytischen Zugriffe werden verschiedene methodische Ansätze – Auswahlkriterien, Datenarten, Analyseprozesse – ebenso beleuchtet wie Herausforderungen und Grenzen der Indikatoren.

Ergebnisse

Die Analyse und Kombination repräsentativer Daten, die regelmäßig und somit in einer Zeitreihe ausgewertet werden können, ermöglicht die Konstruktion von Indikatoren, mit denen sich periodenübergreifender Trends identifizieren lassen.

Erste Ergebnisse zu einzelnen Indikatoren zeichnen ein gemischtes Bild: So zeigt eine systematische Analyse der seit 2014 erfassten politikdidaktischen Promotionen zwar eine kontinuierliche Qualifikation von wissenschaftlichen Nachwuchs, aber auch deutliche Disparitäten hinsichtlich der Repräsentation unterschiedlicher Regionen in Deutschland, hinsichtlich der methodischen Zugriffe und hinsichtlich der untersuchen Bildungsbereiche. Ein weiterer Indikator, der auf der inhaltsanalytischen Auswertung von Modulhandbüchern an 41 repräsentativ ausgewählten Hochschulen beruht, ermöglicht Aussagen zur Verankerung politischer Bildung im bildungswissenschaftlichen Anteil des allgemeinen Lehramtsstudiums Es zeigt sich, dass in 20% der Studiengänge Hinweise auf einschlägige Inhalte vollständig fehlen. Bei 80% finden sich (zumeist minimale) Anteile politischer Bildung. Dabei handelt es sich insbesondere um politische Bildung im Sinne eines sozialen und gesellschaftlichen Lernens. Kernbegriffe einer politischen Bildung im engeren Sinn wie „Grundgesetz“ oder „Rechtstaat“ fanden sich in den Modulhandbüchern nicht. Der dritte ausgewählte Indikator basiert auf der amtlichen Hochschulstatistik und zeigt, wie die Studierenden- und Absolventenzahlen im einschlägigen Fachgebiert unterhalb der Schwelle des Bedarfs und entkoppelt von der allgemeinen Entwicklung der Studierendenzahlen stagnieren.

Indikatoren für die außerschulische politische Bildung: Chancen und Herausforderungen eines Monitorings am Beispiel der Volkshochschulen

Valeriia Hulkovych, Monika Oberle
Georg-August-Universität Göttingen

Theoretischer Hintergrund

In einer von vielfältigen Krisen und zunehmender gesellschaftlicher Polarisierung geprägten Zeit wachsen die Erwartungen an die politische Bildung zur Förderung demokratischer Resilienz und Beteiligung. Um die Potenziale der außerschulischen politischen Bildung einschätzen und effektiv stärken zu können, bedarf es datengestützter Bestandsaufnahmen und Analysen. Ein belastbares,

wiederkehrendes Monitoring, das Aufschluss über Angebote und Rahmenbedingungen, Prozessqualitäten und Wirkungen des Feldes geben und damit eine Basis für konkrete Handlungsempfehlungen legen könnte, ist jedoch bislang ein Desiderat: Die wenigen bisher unternommenen Versuche beschränken sich auf Momentaufnahmen bzw. sind wenig datengesättigt (vgl. z.B. Lange 2010; Kalina 2014; Becker 2019; BMFSFJ 2020). Besondere Herausforderungen für die Erfassung des Feldes liegen dabei einerseits in der Vielzahl variierender Konzepte politischer Bildung und andererseits in einer Vielzahl an Akteuren, die unterschiedliche strukturelle Angebundenheiten und Voraussetzungen aufweisen, wozu auch das föderale Mehrebenensystem öffentlicher Förderung beiträgt. Hieraus ergibt sich nicht nur die Notwendigkeit einer differenzierten und transparenten Eingrenzung des Untersuchungsgegenstandes, sondern auch die Schwierigkeit der Erschließung, Aufbereitung und Harmonisierung von Daten aus zahlreichen Quellen.

Im hier vorgestellten Monitoring-Projekt wurden elf Gruppen von Kernakteuren der außerschulischen politischen Bildung identifiziert, zu denen relevante Daten gesammelt, aufbereitet und analysiert werden. Einer der ausgewählten Akteure – in der Fläche agierend, quantitativ bedeutsam und für sein besonderes Potenzial der Erreichung relevanter Zielgruppen geschätzt – ist die Volkshochschule (Hufer 2016; Martin/Reichart 2020; Scheidig 2022).

Vor diesem Hintergrund sollen Ansätze und Ergebnisse eines Monitorings außerschulischer politischer Bildung beispielhaft am Akteur Volkshochschule diskutiert werden. Die Betrachtung erfolgt entlang der Fragestellung: Mit welchen Indikatoren lässt sich die politische Bildungstätigkeit der Volkshochschulen möglichst umfassend, repräsentativ und wiederkehrend darstellen?

Methode

Trotz des – in vielen Bundesländern sogar gesetzlich festgeschriebenen – Grundauftrags politischer Bildung der Volkshochschulen erlauben die Zuschnitte ihrer sieben Programmbereiche zunächst keine präzise Extrahierung der Aktivitäten politischer Bildung. Um Inhalte politischer Bildung im Angebot der Volkshochschulen möglichst trennscharf identifizieren und erfassen zu können, wurden daher in einem ersten Schritt Expertengespräche mit Volkshochschulvertreter:innen der kommunalen, der Landes- und der Bundesebene geführt. Neben der Eingrenzung des Untersuchungsgegenstandes wurde dabei das Ziel verfolgt, bei der VHS vorliegende, jedoch nicht öffentlich verfügbare Daten zu eruieren. Auf kommunaler Ebene sind dies Daten der einzelnen Einrichtungen, auf Länderebene Daten, die die Landesverbände bereithalten, und auf Bundesebene Daten, die am Deutschen Institut für Erwachsenenbildung (DIE) vorliegen.

In diesem Beitrag werden zwei aus vorliegenden Daten zusammengesetzte Indikatoren genauer vorgestellt und dabei auf methodische Vorgehensweisen, Analyseergebnisse, Herausforderungen und Limitationen eingegangen: Der erste Indikator „Dichte der volkshochschulischen politischen Bildung nach Region“ wurde aus Daten der für die Bundesebene vorliegenden DIE-Volkshochschulstatistik zusammengesetzt. Berechnet wurde die Entwicklung der Veranstaltungsstunden und Teilnehmendenzahlen für ausgewählte Themengebiete des Programmbereichs 1 Politik/Gesellschaft/Umwelt (BAK Politik/Gesellschaft/Umwelt 2015) seit dem Jahr 2010 je Region pro tausend Einwohner:innen. Für den zweiten Indikator „Anzahl der

Veranstaltungen politischer Bildung nach Zielgruppen“ wurden zusätzlich zu den Daten der VHS-Statistik (Alter und Geschlecht der Teilnehmenden klassischer politischer Bildungsveranstaltungen) Daten des Landesverbandes Bayern ausgewertet. Dabei wurde das Veranstaltungsprogramm der bayrischen Volkshochschulen für den Zeitraum 2010 bis 2022 darauf untersucht, inwieweit sich Inhalte politischer Bildung auch im Rahmen von Orientierungs-, Grundbildungs- Integrations- und Schulabschlusskursen finden. Diese Sonderauswertung erlaubt Rückschlüsse bezüglich der Angebote für Zielgruppen, die mit klassischen außerschulischen politischen Bildungsformaten nur selten erreicht werden.

Ergebnisse

Anhand der skizzierten Indikatoren wird die Möglichkeit der Analyse besonderer Potentiale von Volkshochschulen für die politische Bildung kritisch diskutiert: Die Dichte volkshochschulischer politischer Bildung je Region gibt Aufschluss darüber, inwiefern die Existenz von Volkshochschulen auch zu einem Angebot politischer Bildungsveranstaltungen in der Fläche (z.B. in ländlichen Räumen) führt. Der Umfang politischer Bildung nach Zielgruppen gibt Auskunft über die Adressierung der durch konventionelle politische Bildung schwer zu erschließenden Zielgruppen, über deren tatsächliche Erreichung allerdings datenbasiert nur eingeschränkt Aussagen getätigt werden können.

G-16: Kognitive Aktivierung ohne Adaptivität? Notwendigkeit, Möglichkeiten und Herausforderungen einer Verknüpfung beider Konzepte

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* T03 R03 D89

Kognitive Aktivierung ohne Adaptivität? Notwendigkeit, Möglichkeiten und Herausforderungen einer Verknüpfung beider Konzepte

Chair(s): **Jasmin Decristan** (Universität Wuppertal), **Anna-Katharina Praetorius** (Universität Zürich, Switzerland)
*Diskutant*in(nen):* **Thilo Kleickmann** (Universität Kiel)

Kognitive Aktivierung stellt mittlerweile eines der populärsten Konstrukte der Forschung zu Unterrichtsqualität dar (im Überblick Praetorius et al., 2018). Kognitive Aktivierung zielt darauf ab, „Lernende zum vertieften Nachdenken und zu einer elaborierten Auseinandersetzung mit dem Unterrichtsgegenstand“ (Lipowsky, 2020, S. 92) anzuregen. Dabei wird Lernen entsprechend eines konstruktivistischen Verständnisses als ein intrapsychischer und interindividuell unterschiedlicher Konstruktionsprozess angesehen. Vor diesem Hintergrund gilt es, kognitiv aktivierenden Unterricht passend zu den Lernvoraussetzungen zu gestalten und im Unterricht an das jeweilige Vorwissen und die Erfahrungswelt der Schüler*innen anzuknüpfen (Klieme, 2019).

Damit trifft kognitive Aktivierung den Kern von adaptivem Unterricht – verstanden als Passung von unterrichtlichem Handeln der Lehrkräfte zu den individuellen Bedarfen der Lernenden (Corno, 2008; Corno & Snow, 1986). Adaptivität ist fester Bestandteil der Forschungsliteratur (zusammenfassend Parsons et al., 2018) und hat nicht zuletzt durch die erneute Debatte zu individueller Förderung große Aufmerksamkeit erfahren (Dumont, 2019).

Das Verhältnis von kognitiver Aktivierung und Adaptivität ist in der bisherigen Forschung jedoch weder theoretisch-konzeptionell noch empirisch hinreichend beleuchtet worden. Adaptivität wurde teils als für alle Dimensionen von Unterrichtsqualität relevantes Meta-Prinzip von Lehren und Lernen begriffen (Charalambous & Praetorius, 2020; Helmke, 2006), teils aber auch als eigenständiges Merkmal in Listen zu effektivem Unterricht angeführt (Walberg & Paik, 2000).

Das Symposium zielt entsprechend auf eine kritische Diskussion und konstruktive Weiterentwicklung der Konzeptualisierungen und Operationalisierungen von kognitiver Aktivierung und Adaptivität ab. Die vier hierfür zusammengestellten Beiträge des Symposiums arbeiten Gemeinsamkeiten und Unterschiede auf theoretisch-konzeptueller Ebene heraus und stellen entsprechende Operationalisierungen und empirische Befunde vor. Der erste Beitrag vertieft das theoretische Zusammenspiel der beiden Zugänge und leitet zu entsprechenden forschungsbezogenen und praktischen Implikationen über. Der zweite Beitrag wiederum prüft das interaktive Zusammenspiel von Adaptivität und Kognitiver Aktivierung empirisch anhand von Daten aus acht verschiedenen Ländern. Schließlich widmen sich die beiden weiteren Beiträge der Operationalisierung beider Konstrukte. So fokussiert der dritte Beitrag die kognitive Aktivierung in Schüler*innen-Befragungen, trennt zwischen Item-Formulierungen mit und ohne Bezüge zu Adaptivität und prüft jeweilige Zusammenhänge mit Leseleistungen. Der vierte Beitrag wiederum widmet sich der Herausforderung einer angemessenen empirischen Erfassung von Adaptivität und stellt entsprechende Befunde anhand von tablet-basierten Experience Samplings vor.

Beiträge des Symposiums

Konzeptionelle Überlegungen zur Verbindung von adaptivem Unterricht und kognitiver Aktivierung

Benjamin Fauth[1], Jasmin Decristan[2], Hanna Dumont[3]
[1]IBBW, [2]Universität Wuppertal, [3]Universität Potsdam

In diesem konzeptionellen Beitrag werden zwei Konzepte, die zurzeit in der Unterrichtsforschung intensiv diskutiert werden, miteinander in Beziehung gesetzt: Adaptiver Unterricht und kognitive Aktivierung. Es wird untersucht, inwiefern eine stärkere Verzahnung beider Konzepte für die Unterrichtsforschung und für die Weiterentwicklung der unterrichtlichen Praxis hilfreich wäre.

Theoretischer Hintergrund

Bislang hat es erst wenige Ansätze gegeben, die adaptiven Unterricht (z.B. Corno, 2008) und kognitive Aktivierung (z.B. Baumert et al., 2010; Klieme et al. 2009) explizit miteinander in Beziehung setzen. Pietsch et al. (2009) haben eine hierarchische Anordnung von Unterrichtsqualitätsdimensionen vorgeschlagen, wobei Merkmale wie Individualisierung und Differenzierung – als zentrale Merkmale von adaptivem Unterricht – an der Spitze stehen. Hier werden Merkmale adaptiven Unterrichts also als eine eigene Dimension spezifiziert. Im Gegensatz dazu gehen Dumont (2019) und auch Charalambous und Praetorius (2020) davon aus, dass Adaptivität gewissermaßen quer zu anderen Dimensionen der Unterrichtsqualität liegt und damit jeweils ein Bestandteil der einzelnen Dimensionen ist.

Fragestellungen

Wo liegen konzeptionelle Überschneidungen und Unterschiede zwischen den Konzepten der kognitiven Aktivierung und des adaptiven Unterrichts? Welche Implikationen hätte eine stärkere Verzahnung beider Konzepte für die empirische Unterrichtsforschung und für die Weiterentwicklung von Unterricht?

Zur Beantwortung dieser Fragen wird eine systematische Literaturrecherche durchgeführt. Ein besonderer Fokus wird dabei auf die Bedeutung der Konzepte für die individuelle Lernentwicklung und auf deren Implikationen für die Unterrichtsentwicklung gelegt.

Ergebnisse und ihre Bedeutung

In einer ersten Literaturrecherche deutet sich an, dass die individuelle Perspektive in der Literatur zur kognitiven Aktivierung weniger präsent ist, als es die theoretische Fundierung des Konzepts in kognitiv-konstruktivistischen Lerntheorien zunächst vermuten lassen würde (z.B. Fauth et al., 2014; Lipowsky et al., 2009). Dennoch ist es ein Kernmerkmal sowohl des Konzepts der kognitiven Aktivierung als auch des adaptiven Unterrichts, dass die individuelle Lernentwicklung von Schülerinnen und Schülern im Mittelpunkt steht. Dabei fokussiert die kognitive Aktivierung stark auf die (kognitive) Entwicklung des konzeptuellen Verständnisses, während in der Literatur zum adaptiven Unterricht ein breiteres Verständnis zugrunde gelegt wird, indem daneben auch motivationale Aspekte, Persönlichkeit und das selbstregulierte Lernen eine wichtige Rolle spielen (Dumont, 2019).

Von einer stärkeren Verbindung beider Konzepte könnten sowohl die Unterrichtsforschung als auch die Unterrichtspraxis profitieren. In der Unterrichtsforschung wird kognitive Aktivierung vor allem über hoch-inferente Videoanalysen, über klassenaggregierte Urteile von Schülerinnen und Schülern, oder

über die Einschätzung von Aufgabenmaterial erfasst. In all diesen Ansätzen wird jedoch der individuelle Schüler und die individuelle Schülerin nicht hinreichend berücksichtigt, um kognitive Aktivierung auch differenziell untersuchen zu können. Hier könnte die Unterrichtsqualitätsforschung von Erhebungsverfahren profitieren, die im Kontext von Forschung zum adaptiven Unterricht entwickelt wurden und stärker auf die individuelle Passung zwischen Lernvoraussetzungen und Unterrichtsangebot fokussieren. In Frage kommen z.B. Experience Sampling-Verfahren (Dumont, 2019 oder Videokodierungen auf der Individualebene (Decristan et al., 2019). Aus Sicht der Forschung zum adaptiven Unterricht könnten die Prinzipien der kognitiven Aktivierung helfen, um die teilweise geringen Effekte von individualisierten Lernsettings zu erklären (Hattie, 2009).

Hinsichtlich des Potenzials beider Konzepte für die Weiterentwicklung der unterrichtlichen Praxis zeigt sich, dass ihre Umsetzung Lehrkräfte vor ganz ähnliche Herausforderungen stellt: Lehrkräfte benötigen kontinuierliche Informationen zu den Lernverläufen ihrer Schülerinnen und Schüler (Fauth & Leuders, 2021). Es liegt nahe, dass hier neuere Entwicklungen im Bereich der computerbasierten Lernverlaufsdiagnostik und des formativen Assessments eine Lösung sein können - wenn sie in der Lage sind, Lehrkräfte mit den richtigen diagnostischen Informationen und mit Hinweisen zur effektiven Förderung zu versorgen (Souvignier et al., 2020). Auch hier erinnern die Prinzipien der kognitiven Aktivierung daran, dass der Versuch, Unterricht individualisiert zu gestalten, weder mit einer kognitiven Überforderung noch mit einer Unterforderung („Schonung“) gerade von leistungsschwächeren Schülerinnen und Schülern einhergehen sollte.

Untersuchung des Zusammenspiels von kognitiver Aktivierung und Adaptivität durch Interaktionsterme: Eine hinreichende Vorgehensweise?

Anna-Katharina Praetorius[1], Charalambos Charalambous[2]
[1]Universität Zürich, [2]University of Cyprus

Theoretischer Hintergrund

Basierend auf der Argumentation, dass eine optimale Handhabung des komplexen Phänomens Unterrichtsqualität für die Unterrichtsforschung darin besteht, verschiedene Konstrukte zu unterscheiden und diese separat zu untersuchen (z.B. Klieme, 2019), werden kognitive Aktivierung und Adaptivität in bisherigen Studien oft über separate Skalen erfasst und in Analysen einbezogen (siehe z.B. die PISA-Erhebungen). Aktuelle theoretische Überlegungen hingegen betonen die enge Verzahnung von kognitiver Aktivierung und Adaptivität und dass die beiden Konstrukte entsprechend nicht unabhängig voneinander gedacht werden können (Charalambous & Praetorius, 2020). Folgt man dieser Überlegung, stellt sich die Frage, wie dieses Zusammenspiel empirisch untersucht werden kann. Soll dies im Anschluss an die bisherige Forschungspraxis erfolgen und entsprechend separate Skalen einbezogen werden, lassen sich statistisch zwei mögliche Herangehensweisen unterscheiden: (a) eine sequentielle Zusammenführung, bei der in den Blick genommen wird, wie sich das eine Konstrukt auf das andere auswirkt (Mediation) und (b) die Untersuchung, wie das eine Konstrukt die Funktionsweise des anderen verändern kann (Moderation). Während (a) konzeptuell bei den beiden Konstrukten wenig überzeugend scheint, stellt (b) eine sinnvolle Möglichkeit einer Überprüfung des Zusammenspiels von kognitiver Aktivierung und Adaptivität dar.

Fragestellungen

1) Erhöht sich die prädiktive Validität in Bezug auf die Entwicklung von Schülerleistungen sowie Schülermotivation, wenn die Interaktion zwischen kognitiver Aktivierung und Adaptivität explizit in die Datenanalysen einbezogen wird?

2) Wie konsistent sind diese Effekte über verschiedene Länder hinweg?

Methode

Im Rahmen der TALIS-Videostudie (OECD, 2020), die auf die Untersuchung von Unterrichtseinheiten zum Thema quadratische Gleichungen fokussiert, wurden Daten von insgesamt ca. 700 Lehrpersonen und 17 500 Schüler*innen der Sekundarstufe aus acht Ländern gesammelt. Für die Zwecke der vorliegenden Arbeit wurden für jedes Land separate Mehrebenenmodelle (Schüler*innen innerhalb von Klassen) mit drei abhängigen Variablen (die Testergebnisse in Bezug auf quadratische Gleichungen, das Interesse der Schüler*innen an Mathematik und die Selbstwirksamkeit in Mathematik) durchgeführt und für verschiedene Kovariaten auf Schüler*innen- und Klassenebene (Geschlecht der Schüler*innen, SES, Migrationshintergrund) sowie die kognitiven oder affektiven Masse vor der Unterrichtseinheit (allgemeiner Leistungstest, Interesse und Selbstwirksamkeit in Mathematik im vorangegangenen Jahr) kontrolliert. Kognitive Aktivierung (4 Items) und Adaptivität (5 Items) wurden aus Sicht der Schüler*innen erfasst. Die Analysen wurden zunächst für beide Konstrukte separat durchgeführt, in einem nächsten Schritt wurden beide Konstrukte parallel in das Modell aufgenommen und abschliessend auch deren Interaktion.

Ergebnisse und ihre Bedeutung

Betrachtet man die Leistungsentwicklung, lassen sich sowohl auf Klassen- als auch auf Individualebene einige signifikante Haupteffekte für kognitive Aktivierung und/oder Adaptivität identifizieren, die jedoch nicht konsistent über Länder hinweg sind. Interaktionseffekte konnten in keinem der Länder identifiziert werden. Bezüglich der beiden affektiven Outcomes konnten auf beiden Ebenen und für beide Dimensionen in etlichen Ländern positive Haupteffekte gefunden werden, die zum Teil auch bei paralleler Aufnahme der Dimensionen in das Modell erhalten blieben. Interaktionseffekte konnten auf Ebene 1 in mehreren Ländern nicht identifiziert werden (Chile, Mexiko, Kolumbien), in einigen Ländern (England, Deutschland) fielen diese positiv und in anderen Ländern (China und Japan) negativ aus; in keinem Fall konnte dieser Effekt durchgängig bei beiden affektiven Outcomes identifiziert werden. Auf Ebene 2 konnten keinerlei Interaktionseffekte identifiziert werden.

Die gefundenen Moderationseffekte könnte man dahin gehend interpretieren, dass eine Betrachtung des Zusammenspiels von kognitiver Aktivierung und Adaptivität sinnvoll erscheint und dieses Zusammenspiel in Zusammenhang mit bestimmten Ländermerkmalen zu stehen scheint. Kritisch betrachtet deutet die insgesamt geringe Anzahl an Moderationseffekten darauf hin, dass die Integration von kognitiver Aktivierung und Adaptivität von Beginn einer Studie an – von den Phasen der Konzeptualisierung und Operationalisierung bis hin zur Messung – gedacht werden müsste (für eine ähnliche Forderung für die Unterrichtsforschung, siehe Hiebert & Stigler, in Druck).

Kognitive Aktivierung in Befragungen von Schülerinnen und Schülern – Unterscheidung zwischen dem Potential zur kognitiven Aktivierung und individueller kognitiver Aktivierung

Svenja Rieser, Jasmin Decristan
Universität Wuppertal

Theoretischer Hintergrund

Ein kognitiv aktivierender Unterricht gilt als Unterrichtqualitätsmerkmal und soll vertiefte kognitive Verarbeitungsprozesse bei den Lernenden auslösen und dadurch entscheidend zum Lernerfolg beitragen (z.B. Klieme, 2018). Die empirischen Befunde zum Zusammenhang zwischen kognitiver Aktivierung und dem Lernerfolg sind jedoch bislang nicht eindeutig (Praetorius et al., 2018). Eine mögliche Ursache dafür könnte in unterschiedlichen Konzeptualisierungen von kognitiver Aktivierung liegen.

Üblicherweise werden bei kognitiver Aktivierung Handlungen und Aufgabenstellungen der Lehrkraft fokussiert, von denen angenommen wird, dass sie eine vertiefte Verarbeitung bei den Lernenden anregen können. Dieses Klassenmerkmal wird auch als „Potential zur kognitiven Aktivierung“ (PKA; Kunter & Voss, 2011) beschrieben.

Die alleinige Fokussierung auf das PKA wird der kognitiven Aktivierung jedoch nicht gerecht (Herbert & Schweig, 2020). Mit Bezug auf konstruktivistische Annahmen zum Lernen als individueller und aktiver Konstruktionsprozess (Palinscar, 1998; Webb, 2011), lässt sich annehmen, dass ein identisches Lehrkraftverhalten von den Lernenden, abhängig von ihren jeweiligen Lernvoraussetzungen, als unterschiedlich kognitiv aktivierend erlebt wird. Entsprechend können Lernende innerhalb einer Klasse trotz gleichem PKA eine unterschiedliche „individuelle kognitive Aktivierung“ (IKA) erleben.

Beide Konzeptualisierungen von kognitiver Aktivierung lassen sich auch in Items zur Befragung von Lernenden identifizieren. Ziel dieses Beitrags ist es daher folgende Fragestellungen zu untersuchen:

1. Lassen sich Items zur Erfassung des Potentials zur kognitiven Aktivierung oder der individuellen kognitiven Aktivierung in Befragungen von Lernenden innerhalb und zwischen Klassen empirisch trennen?

2. Welche Zusammenhänge bestehen zwischen dem Potential zur kognitiven Aktivierung bzw. der individuellen kognitiven Aktivierung und dem Lernerfolg innerhalb und zwischen Klassen?

Methode

Die Analysen basieren auf Angaben von 3797 Lernenden aus 192 vierten Klassen (IGLU 2016; Hußmann et al., 2020). Die Kinder beantworteten zehn Items zur kognitiven Aktivierung im Deutschunterricht, von denen sechs dem PKA (z.B. „Unsere Deutschlehrer*in möchte, dass wir unsere Antworten erklären.“) und vier der IKA (z.B. „Unsere Deutschlehrer*in stellt Aufgaben, die schwierig sind.“) zugeordnet wurden. Die Leseleistung wurde durch den standardisierten Lesekompetenztest aus IGLU 2016 erfasst.

Zur Beantwortung der ersten Fragestellung wurde mittels Mehrebenen-konfirmatorischer Faktorenanalysen (ML-CFA) geprüft, ob ein Modell mit zwei Faktoren (PKA und IKA) auf Individual- und Klassenebene (2-2-Modell) besser zu den Daten passt als alternative Modelle.

Zur Beantwortung der zweiten Fragestellung wurden doppelt latente Mehrebenen-Regressionsanalysen herangezogen. Dabei wurde, unter Einbezug von Kovariaten, der

Zusammenhang zwischen den beiden Facetten der kognitiven Aktivierung und der Leseleistung geprüft.

Ergebnisse und ihre Bedeutung

Die Ergebnisse der ML-CFA Modelle zeigen, dass es sich beim PKA und der IKA nicht nur theoretisch, sondern auch empirisch um zwei voneinander abgrenzbare Konstrukte handelt.

In den Mehrebenen-Regressionsanalysen konnte für das PKA nur auf Klassenebene ein positiver Zusammenhang mit der Leseleistung gefunden werden (β = .17*; SE = 0.08). Somit stützen unsere Ergebnisse die konzeptionelle Verortung des PKA als Unterrichtsqualitätsdimension und damit als zwischen Klassen variierendes Merkmal (vgl. Marsh et al., 2012).

Für die IKA fand sich hingegen nur ein negativer Zusammenhang mit der Leseleistung auf Individualebene (β = -.25**; SE = 0.03). Ergänzende Analysen unterstützen die Überlegung, dass die Items zur IKA auch ein Indikator der Überforderung statt der erlebten Herausforderung darstellen könnten. Weiterhin könnte der Unterricht an den leistungsschwächeren Kindern ausgerichtet sein, so dass leistungsstarke Kinder vermutlich eine geringe kognitive Aktivierung berichten, aber trotzdem eine hohe Leseleistungen zeigen.

Die Ergebnisse tragen somit einerseits zur Erklärung der uneindeutigen Befunde zum Zusammenhang von kognitiver Aktivierung und dem Lernerfolg bei. Andererseits liefern sie weitere Diskussionspunkte zum Verhältnis von kognitiver Aktivierung und Adaptivität. Wie kann kognitive Aktivierung unter Einbezug konstruktivistischer Annahmen von Lehren und Lernen konzeptualisiert und operationalisiert werden? Wäre eine hohe Übereinstimmung der Lernenden innerhalb von Klassen bei Items zur IKA ein Indikator für Adaptivität? Vor diesem Hintergrund werden im Beitrag weitere Implikationen für das Verhältnis von kognitiver Aktivierung und Adaptivität eingehender beleuchtet.

Subjektive Schwierigkeitseinschätzungen als Zugang zur Erfassung von Adaptivität

Hanna Dumont, Simon Ohl
Universität Potsdam

Ziel von adaptivem Unterricht ist es, das Unterrichtsangebot an die individuellen Lernvoraussetzungen der einzelnen Schüler*innen anzupassen, um jeden Einzelnen optimal zu fördern (Corno, 2008; Hardy, Decristan & Klieme, 2019). Basierend auf den Erkenntnissen der Forschung zu Aptitude-Treatment-Interactions wird dazu der Grad an Unterstützung durch die Lehrperson mit zunehmendem Kompetenzniveau der Lernenden reduziert. Dabei gilt es, die Zone der proximalen Entwicklung (Vygotsky, 1978) für jeden zu finden. Adaptivität bedeutet somit eine kognitive Überlastung als auch „Unterlastung“ und im Sinne der Cognitive Load Theory (Sweller, van Merrienboer & Paas, 1998) zu vermeiden. Doch wie lässt sich die Adaptivität zwischen Unterrichtsangebot der Lehrperson und Lernvoraussetzung der Schüler*innen – im Sinne eines Merkmals der Tiefenstruktur von Unterricht – empirisch erfassen? Es ist davon auszugehen, dass letztlich nur die Schüler*innen selber darüber Auskunft geben können, ob sie kognitiv unter- oder überfordert sind. Diese Annahme wird durch die Forschung zur Cognitive Load Theory gestützt, in der subjektive Einschätzungen zur mentalen Anstrengung und zur wahrgenommenen Schwierigkeit das

verbreitetste Maß zur Erfassung von kognitiver Anstrengung darstellen (Paas, Tuovinen, Tabbers & Van Gerven, 2003) und somit auf den Schulkontext übertragen werden können. Dabei muss jedoch bedacht werden, dass subjektive Schwierigkeitseinschätzungen von Schüler*innen von ihrem akademischen Selbstkonzept beeinflusst sein könnten. Um solche Verzerrungen zu minimieren, könnte eine Möglichkeit darin bestehen, Schüler*innen mehrfach und möglichst unmittelbar nach dem Unterricht im Sinne der Experience-Sampling-Methode (Zirkel, Garcia & Murphy, 2015) zu befragen.

Vor diesem Hintergrund geht der vorliegende Beitrag der Frage nach, inwiefern sich subjektive Schwierigkeitseinschätzungen des Unterrichts durch Schüler*innen, die im Sinne der Experience-Sampling-Methode unmittelbar nach einer Unterrichtseinheit erhoben werden, zur Erfassung des Grads der Adaptivität eignen. Dazu wurde eine tablet-basierte App mit insgesamt neun Fragen zum Unterrichtserleben entwickelt. Neben der im Fokus stehenden wahrgenommen Schwierigkeit der vorangegangenen Unterrichtseinheit (erfasst über eine kontinuierliche Skala von 0 = viel zu leicht über 50 = genau richtig bis 100 = viel zu schwer) wurde nach der Konzentrationsfähigkeit, der Anstrengung, dem Grad der kognitiven Aktivierung, dem Interesse, dem Wohlbefinden sowie dem subjektiven Nutzen (jeweils über eine 5-stufige Likert-Skala) gefragt. Pilotiert wurde die App auf der Basis einer Stichprobe von 57 Grundschüler*innen aus 4 Klassen sowie 94 Sekundarschüler*innen aus 7 Klassen jeweils einer Schule, wobei die App für eine Schulwoche nach jeder Unterrichtsstunde zum Einsatz kam. Anschließend wurde die App im Rahmen einer größeren Studie an 23 Klassen aus 6 Grundschulen über eine Schulwoche hinweg eingesetzt, so dass insgesamt Daten von 457 Schüler*innen und 262 Unterrichtseinheiten vorliegen.

Zur Erfassung des Grads an Adaptivität wurden die Antworten auf die Frage zur wahrgenommenen Schwierigkeit so transformiert, dass die Abweichungen vom theoretischen Optimum 50 = genau richtig berechnet wurden, d.h. je niedriger der Wert desto höher die Adaptivität. Im Anschluss daran erfolgte eine Varianzzerlegung in Anlehnung an Parrisius, Gaspard, Zitzmann, Trautwein und Nagengast (2022). Diese ergab, dass insgesamt 94% der Varianz auf der Individualebene lag, wobei 60% auf Unterschiede zwischen Unterrichtseinheiten (Intraindividuelle Ebene) und 34% auf ein konsistentes Antwortverhalten über Unterrichtseinheiten hinweg (Interindividuelle Ebene) zurückzuführen war. Die 6% Varianz auf Gruppenebene setzten sich zu 3% aus Varianz zwischen Unterrichtseinheiten (Intragruppenebene) und zu 3% aus über Unterrichtseinheiten geteilter Varianz (Intergruppenebene) zusammen. Des Weiteren zeigten sich auf allen Ebenen theoretisch plausible Zusammenhänge zwischen dem Grad der Adaptivität von Unterrichtseinheiten und den übrigen Fragen zum Unterrichtserleben, wie beispielsweise dem Grad der kognitiven Aktivierung (Intraindividuelle Ebene: r = -.15, Interindividuelle Ebene: r = -.16, Intragruppenebene: r = -.28, Intergruppenebene: r = -.13). Zusammenfassend deuten die Ergebnisse darauf hin, dass sich mittels Experience Sampling gewonnene subjektive Schwierigkeitseinschätzungen zur Erfassung von Adaptivität eignen könnten und Adaptivität mit kognitiver Aktivierung einhergeht.

G-17: Förderung von Lesekompetenz – What works?

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* R11 T00 D01

Förderung von Lesekompetenz – What works?

Chair(s): **Annika Ohle-Peters** (TU Dortmund, Germany), **Nele McElvany** (TU Dortmund, Germany)
*Diskutant*in(nen):* **Nadine Spörer** (Universität Potsdam)

Lesekompetenz ist eine entscheidende Voraussetzung für schulisches Lernen in allen Domänen, für Bildungserfolg und für gesellschaftliche Teilhabe (Artelt & Dörfler, 2010). Vor diesem Hintergrund ist die Förderung von Lesekompetenz und deren Vorläuferfähigkeiten eine wichtige Aufgabe des Grundschulunterrichts. Theoretisch fundierte und empirisch mit adäquaten Untersuchungsdesigns evaluierte Maßnahmen zur Förderung von Wortschatz, Lesestrategien, -flüssigkeit, - motivation und -verständnis sind sowohl für Bildungsforschung als auch -praxis von großer Relevanz. In den vier Studien dieses Symposiums werden systematisch die Wirkungen verschiedener Leseförderinterventionen und deren Implementationsbedingungen empirisch überprüft und hinsichtlich theoretischer und praktischer Implikationen diskutiert.

Im ersten Beitrag wurden in einer quasi-experimentellen Studie Wirksamkeit und Wirkweise einer peer-gestützten Leseförderung im regulären Unterricht der dritten Klassenstufe untersucht. Im Fokus der Intervention mit Prä-Post-Follow-Up-Test Design stand die Vermittlung von Lesestrategien, die in den Analysen als Mediator für die Entwicklung von Lesemotivation, -flüssigkeit und Wortschatz identifizieren ließen.

In Beitrag zwei wurden in einem randomisierten Experimentaldesign verschiedene Ansätze einer tabletbasierten Wortschatzförderung hinsichtlich ihrer Wirksamkeit in der vierten Klassenstufe untersucht. Verglichen wurden implizite Verfahren mit zusätzlichen schriftlichen, auditiven oder bildlichen Erklärungen des neuen Wortschatzes mit einer expliziten Wortschatzförderung und einer Kontrollgruppe ohne neuen Wortschatz. Die Ergebnisse zeigten einen Wortschatzzugewinn in den impliziten und expliziten Förderbedingungen; Unterschiede zwischen schriftlichen, auditiven und bildlichen Erklärungen konnten nicht identifiziert werden.

Beitrag drei untersucht mit einem Prä-Post-Kontrollgruppendesign Effekte von persönlichen Voraussetzungen sowie Implementations- und Prozessmerkmalen auf die Effekte einer Intervention zur Förderung von Leseflüssigkeit durch Paired Reading mit Eltern und Lesecoaches in der dritten Klasse. Die Befunde weisen auf signifikante Unterschiede zwischen Eltern und Lesecoaches hinsichtlich der Prozessvariablen hin, differenzielle Effekte auf die Leseflüssigkeit der Lernenden konnten jedoch nicht gezeigt werden.

Der vierte Beitrag fokussiert auf die Implementation von Lesefördermaßnahmen mit dem Fokus auf die Förderintensität. In dieser Studie wurden Lehramtsstudierende für das Leseförderprogramm „Der Lese-Sportler“, welches auf genaues, flüssiges und verstehendes Lesen abzielt, geschult und setzten dies im Rahmen ihres Praxissemesters über mehrere Monate in den Klassenstufen 2-4 ein. Erste Analysen zeigten, dass die über den regulären Unterricht hinausgehende Förderung mit geringer Intensität implementiert werden konnte.

Die Ergebnisse der vier Studien liefern wertvolle Hinweise für die effektive Förderung von Lesekompetenz. Durch die methodischen Ansätze der Studien können sowohl Rückschlüsse für die

Forschung, beispielsweise zur Lesekompetenzentwicklung, als auch für die schulische Praxis hinsichtlich der Implementation konkreter Fördermaßnahmen gezogen werden. Mit dem Fokus auf die Grundschule fokussiert das Symposium zudem auf eine wichtige Schnittstelle in der Bildungsbiographie von Lernenden.

Beiträge des Symposiums

Prüfung der Wirksamkeit und Wirkweise eine peer-gestützten strategiebasierten Leseförderung in der 3. Klassenstufe

Vanessa Völlinger[1], Lubbe Dirk[2]
[1]Justus-Liebig-Universität Gießen, [2]Medizinische Hochschule Brandenburg Theodor Fontane (MHB)

Nach den ersten Phasen der Leseentwicklung zu Beginn der Schulzeit sollten Kinder bereits in den oberen Grundschulklassen beim Lesen Informationen aus Texten extrahieren, um diese Informationen für ihr Lernen zu nutzen. Schulleistungsstudien wie PIRLS und PISA haben gezeigt, dass sich die Lesekompetenz und insbesondere das Leseverständnis nicht bei allen Schüler:innen im Laufe der Schulzeit bis zu dieser Stufe entwickeln (Mullis et al., 2017; OECD, 2019). Daraus lässt sich der Schluss ziehen, dass Schüler:innen beim sinnentnehmenden Lesen explizit unterstützt werden müssen. Um mit begrenzten Ressourcen möglichst viele Schüler:innen zu erreichen, bedarf es theorie- und evidenzbasierter Trainingskonzepte, die von Lehrkräften auf Klassenebene umgesetzt werden können (Abbott et al., 2011; siehe auch Literatur zum Response-to-Intervention-Modell, evidenzbasierter Unterricht auf Ebene 1). Die Heterogenität der (Lese-)Leistungen von Schüler:innen in ganzen Klassen kann in solchen Programmen aufgefangen und sogar genutzt werden, indem die Schüler:innen gemeinsam an Leseaufgaben arbeiten, um sich gegenseitig in ihrem Leseprozess innerhalb ihrer Zone der proximalen Entwicklung (Vygotksy, 1978) zu unterstützen. Die positiven Auswirkungen der Zusammenarbeit mit anderen Schüler:innen auf die Lernleistung sind durch umfangreiche Forschung zum kooperativen und peer-gestützten Lernen belegt (Kyndt et al., 2013; Rohrbeck et al., 2003; Tenenbaum et al., 2020).

Für den Bereich des Lesens zeigen Metaanalysen und Übersichtsarbeiten die positiven Fördereffekte von peer-gestützten Programme, die strategische Ansätze zum Textverständnis vermitteln auf das Leseverständnis von Schüler:innen (Neitzel et al., 2022; Puzio & Colby, 2013; Slavin et al., 2009). Dementsprechend wurde in der vorliegenden Studie eine peer-gestützte, strategiebasierte Leseintervention auf Klassenebene durchgeführt, um die Umsetzbarkeit und Wirksamkeit eines Trainingsprogramms im regulären Grundschulunterricht zu testen.

Wie von Methodenexperten empfohlen, sollte die Evaluation von Interventionsprogrammen theoriegeleitet sein und nicht nur die Wirksamkeit der Intervention, sondern auch die Entfaltung ihrer Wirkungen berücksichtigen (Chen, 1997). Strukturgleichungsmodelle sind hierfür ein geeigneter Ansatz, da mit diesen die strukturellen Zusammenhänge relevanter Variablen getestet werden können. Diesen Empfehlungen folgend wird in der vorliegenden eine Mediationsmodell getestet, in das die angenommenen Haupteinflussfaktoren des Programms einbezogen werden. Um die längsschnittliche Datenstruktur angemessen zu berücksichtigen und sowohl Unterschiede zwischen den Personen als auch Veränderungen innerhalb der Personen zu untersuchen, wird eine Modellierung simultaner Wachstumskurven verwendet.

Studentische Lesetrainer:innen leiteten im Schuljahr 2019/20 in 11 dritten Klassen eine 20-stündige strategiebasierte und peer-gestützte Intervention im regulären Klassenunterricht an, während 10 weitere Klassen als Wartekontrollgruppe dienten (Teilstudie 1). Im Schuljahr 2021/22 wurde in einer zweiten Studie 18 Trainingsstunden in 10 weiteren Klassen durchgeführt, während wiederum 10 Klassen als Wartekontrollgruppe dienten (Teilstudie 2).

Instruiert und geübt wurde zusätzlich zu flüssigem Lesen mit Peers auch die Anwendung verständnisprüfender und verständnisfördernder Lesestrategien. Die Evaluation der Leistungsentwicklung erfolgte in einem Prä-Post-Follow-Up-Test Design. Dabei wurden in beiden Untersuchungsbedingungen gleichermaßen alle abhängigen Variablen (intrinsische Lesemotivation, Leseflüssigkeit, Wortschatz, Lesestrategien) zu den drei Messzeitpunkten erhoben.

Die primäre Evaluation der Leseförderung wurde anhand des Leseverständnis-Testwertes vorgenommen. In Teilstudie 1 war der Anstieg in der Leseverständnisleistung zum Post-Zeitpunkt bei den geförderten Klassen signifikant größer als in der Wartekontrollgruppe (d = .17). Anhand eines Simultaneous Growth Models zeigte sich, dass ein besserer individueller Ausgangswert im Leseverständnis prädiktiv für eine stärkere Verbesserung über die Messzeitpunkte war. Weiterhin legt das Modell einen mediierenden Effekt von der individuellen Verbesserung der Anwendung von Lesestrategien über die Verbesserung der Leseflüssigkeit zu einem Zuwachs des Leseverständnisses nahe. Die Daten aus Teilstichprobe 2 werden aktuell noch ausgewertet und fließen dann zusätzlich in die Analysen ein.

Die Ergebnisse unterstreichen die Bedeutung verschiedener kognitiver Determinanten für das Leseverständnis und zeigen, dass insbesondere die Lesestrategien als Mediator der Effekte von Lesemotivation, Leseflüssigkeit und Wortschatz eine zentrale Stellung in einem Modell einnehmen. Praktische Implikationen der Studie ergeben sich aus dem Beleg der Umsetzbarkeit der Leseförderung in ganzen Regelschulklassen und der Integration verschiedener Förderelemente in einem kombinierten Programm.

Wie profitieren Grundschulkinder von unterschiedlichen Formaten digitaler Wortschatzförderung?

Annika Ohle-Peters, Erdal Papatga, Nele McElvany
TU Dortmund, IFS

Theorie

Der Wortschatz gilt als Voraussetzung für Leseverständnis und ist damit ein wichtiges Bildungsziel in der Grundschule (u.a. Artuso & Palladino, 2022). Die Frage nach effektiven Wortschatzfördermöglichkeiten ist auch vor dem Hintergrund der großen Heterogenität der Schülerschaft in der Grundschule ein zentraler Bereich der empirischen Bildungsforschung. Für das Lernen mit analogen Medien weisen die bisherige Forschungsergebnisse auf die Wirksamkeit sowohl von impliziten als auch expliziten Förderansätzen hin (Li & Lei, 2022; Ramos, 2015). Damit neue Wortbedeutungen aus dem Kontext erschlossen werden können, sind Leseverständnis und Lesemotivation aufgrund ihres reziproken Verhältnisses von großer Relevanz (Miyamoto, Pfost & Artelt, 2018). Im Elementarbe-reich konnte im digitalen Kontext beim frühen Wortlesen mit EBooks ein zusätzlicher Effekt unterstützender Inputs wie einer Wörterbuchfunktion auf den Wortschatzerwerb gezeigt werden (E-gert,

Cordes & Hartig, 2022). Inwiefern dies auch auf ältere Lernende übertragbar ist und welche Bedeutung zusätzliche Unterstützungsformate sowie Leseverständnis und Lesemotivation für den Wortschatzerwerb mit digitalen Medien haben, war Ziel der vorliegenden Untersuchung.

Forschungsfragen (F)

F1: Lernen Schülerinnen und Schüler (SuS) mehr Wortschatz, wenn sie a) Texte mit bzw. ohne zusätzliche Erklärungen der Zielwörter lesen (implizite Bedingungen mit bzw. ohne Kontext) oder b) ausschließlich schriftliche Erklärungen der Zielwörter erhalten (explizite Bedingung) oder c) Texte mit Synonymen statt Zielwörtern lesen (Kontrollbedingung)?

F2: Unterscheiden sich die Wortschatzzuwächse, wenn SuS a) schriftliche, b) mündliche oder c) bildliche Erklärungen der Zielwörter erhalten?

F3: Gibt es differenzielle Effekte der Wortschatzförderung hinsichtlich a) des Leseverständnisses und b) der Lesemotivation der SuS?

Methode

In der Studie wurden N = 246 Viertklässlerinnen und Viertklässler (MAlter = 10.13 Jahre, SDAlter = 0.75 Jahre, 54.9% Mädchen, 46.4% Migrationshintergrund) innerhalb von Klassen randomisiert auf sechs folgende Untersuchungsbedingungen zugewiesen und erhielten in einer zweistündigen Interven-tion unterschiedliche Fördermaterialien mit unbekanntem Wortschatz (Zielwörter):

A) Zwei Lesetexte + schriftliche Erklärungen der Zielwörter (implizite Bedingung mit zusätzli-cher Unterstützung)

B) Zwei Lesetexte + mündliche Erklärungen der Zielwörter (s.o.)

C) Zwei Lesetexte + bildliche Erklärungen der Zielwörter (s.o.)

D) Schriftliche Erklärungen (explizite Bedingung)

E) Zwei Lesetexte ohne Erklärung (rein implizite Bedingung)

F) Zwei Lesetexte ohne Zielwörter (Kontrollgruppe)

Zu Beginn der Intervention (MZP I) bearbeiteten die SuS unter anderem einen Fragebogen zu Demographie und intrinsischer Lesemotivation, einen Leseverständnistest (ELFE 2; Lenhard, Lenhard & Schneider, 2018) und einen Wortschatztest (McElvany, Ohle, El-Khechen, Hardy & Cinar, 2017), welchen sie jeweils direkt nach dem Lesen eines Textes noch einmal bearbeiteten (MZP II). Die Reliabilität der eingesetzten Instrumente war in allen Gruppen akzeptabel bis gut. Sämtliche Tests, Fragebögen und Interventionsmaterial wurden in Open Sesame programmiert auf Tablets administriert. Zur Beantwortung der Forschungsfragen wurden ANOVA mit Messwiederholung und Migrationshintergrund als Kovariate in F2 aufgrund signifikanter Gruppenunterschiede gerechnet.

Ergebnisse und Implikationen

Für F1 ergab die ANOVA einen signifikanten Haupteffekt des Wortschatzes zu MZP I und einen signifikanten Interaktionseffekt mit der Gruppenzugehörigkeit (A/B/C/E vs. D vs. F), wobei die Kin-der in der expliziten Bedingung am meisten Wortschatz dazu lernten. Bezüglich der Erklärungsmodi (F2) zeigten sich unter Kontrolle des Migrationsstatusses keine signifikanten Interaktionseffekte der

Gruppenzugehörigkeit (A vs. B vs. C). Hinsichtlich der differenziellen Effekte konnten positive Interaktionen sowohl für das Leseverständnis als auch für die Lesemotivation festgestellt werden. Ein Extremgruppenvergleich zeigte in den impliziten Bedingungen nur für die SuS mit hohem Leseverständnis einen Wortschatzzugewinn. Die Befunde dieser Studie weisen auf das Potenzial digita-ler Wortschatzförderung hin. Ein zusätzlicher Effekt von verschiedenen Unterstützungsformaten zu Erklärung unbekannter Wörter konnte in dieser Studie nicht identifiziert werden, dennoch zeigen die Ergebnisse, dass es auch innerhalb einer kurzen Zeit gelungen ist, Wortschatzzugewinne bei den Lernenden nachzuweisen. Die Administration über Tablets ermöglicht eine gut zu realisierende Implementation in den Schulalltag, langfristige Effekte dieser digitalen Wortschatzförderung sollten in zukünftigen Studien untersucht werden.

Paired Reading mit Eltern oder Lesecoachs: Erklären Implementations-, Prozessmerkmale oder persönliche Voraussetzungen der Beteiligten differenzielle Effekte?

Caroline Villiger, Loredana Torchetti
Pädagogische Hochschule Bern

Theoretischer Hintergrund und Fragestellung

Bei Untersuchungen hinsichtlich der Wirksamkeit einer Methode liegt der Fokus oftmals auf Lernergebnissen (Outcomes). Die Frage, wie die Effekte tatsächlich zustande kommen, bleibt dabei weitgehend unbeantwortet. Gerade im Bereich ausserschulischen Lernens sind Lernprozesse und ihre Wirkung noch weitgehend ungeklärt bzw. unerforscht. Dies trifft auch auf die Methode «Paired Reading» (PR), eine Methode zur Förderung der Leseflüssigkeit im dyadischen Setting, zu (vgl. Topping et al., 2012).

Dieser Beitrag basiert auf Daten einer Interventionsstudie bei Schülerinnen und Schülern der dritten Grundschulstufe mit Leseschwierigkeiten. Ein Vergleich von Elterncoachs und freiwilligen Lesecoachs hatte differenzielle Effekte der Methode aufgezeigt (Autor/-innen, 2019): Kinder, die mit einem Lesecoach trainiert hatten, zeigten signifikant mehr Zuwachs in der Leseflüssigkeit als Kinder in der Kontrollgruppe und Kinder, die mit ihren Eltern trainiert hatten. Davon ausgehend zielt der vorliegende Beitrag darauf ab, diese differenziellen Effekte zwischen Eltern und Lesecoachs mittels kausaler Mediation anhand von Implementations- und Prozessdaten (z.B. Befindlichkeit der Beteiligten) aufzuklären. Auch die moderierende Rolle von persönlichen Voraussetzungen der Eltern und Lesecoachs (z.B. Disposition gegenüber dem Lesen) wird geprüft. Als theoretische Grundlage dient das Modell der Inanspruchnahme und Wirkung von extracurricularen Förderangeboten von Lohaus und Wild (2021). Die Forschungsfragen lauten wie folgt: 1. Mediieren Implementations- und Prozessmerkmale den Effekt des PR-Trainings auf die Leseflüssigkeit der Schüler*innen und erklären somit die unterschiedlichen Effekte von Eltern und Lesecoachs? 2. Sind bestimmte Voraussetzungen der Eltern und Lesecoachs verantwortlich für die unterschiedlichen Trainingseffekte?

Methode

Bei der Stichprobe handelt es sich um N = 129 Drittklässlerinnen und Drittklässler, wovon 65 mit einem Elternteil und 64 mit einem freiwilligen Lesecoach trainiert hatten. Die Studie basiert auf einem Prä-Post-Kontrollgruppendesign. Die Leseflüssigkeit dient als abhängige Variable und wurde mit einem standardisierten 1-Minuten-Lesetest erhoben (Walter, 2009). Unter Kontrolle von

Eingangsvoraussetzungen der Kinder werden Implementations- und Prozessvariablen aus unterschiedlichen Quellen (Videodaten, Selbstbericht) in Bezug auf Mediationseffekte und persönliche Voraussetzungen der Eltern bzw. Lesecoachs auf Moderationseffekte hin geprüft. Bei den Variablen handelt es sich um Angaben zur Durchführungsqualität (gemäss vorgängiger Schulung), zum Engagement und zur Befindlichkeit der Beteiligten (Kind und Eltern- bzw. Lesecoach). Weiter wird die Trainingsdauer berücksichtigt. Die Analysen wurden mit den Statistikprogrammen SPSS 28 und MPlus 8.7 (Mediationsanalysen) durchgeführt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen signifikante Unterschiede zwischen Eltern und Lesecoachs hinsichtlich mehrerer Prozessvariablen: Lesecoachs berichten eine bessere emotionale Befindlichkeit während des Trainings, loben das Kind mehr und sprechen länger über den Inhalt des Buches als Eltern, was insgesamt auf eine positive/entspanntere Lernatmosphäre hinweist. Auch bei den Voraussetzungen zeichnet sich ein signifikanter Unterschied ab: Lesecoachs verfügen über eine positivere Disposition gegenüber dem Lesen.

Trotz der nachgewiesenen Unterschiede können die geprüften Variablen die differenziellen Effekte zugunsten der Lesecoachs jedoch nicht erklären – weder Mediations- noch Moderatoreffekte sind statistisch nachweisbar. Dies ist ein Hinweis darauf, wie schwierig es ist, die (bei ausserschulischen Trainingsprogrammen) für die Wirksamkeit der Methode verantwortlichen Komponenten aufzudecken. Für zukünftige Forschungsarbeiten wird empfohlen, mit grösseren Stichproben zu arbeiten, um die Chancen des Nachweises von Wirkmechanismen zu erhöhen.

Immerhin weisen die Ergebnisse zum Vergleich der Implementations- und Prozessdaten auf den unterschiedlichen Charakter der beiden Settings (Training mit Eltern vs. Lesecoachs) hin: In mehrfacher Hinsicht erscheint die Trainingssituation zuhause mit den Eltern als weniger vorteilhaft. Für die Praxis bedeutet dies, dass Kinder mit Leseschwierigkeiten vermutlich aufgrund einer positiveren Lernatmosphäre von der Förderung durch einen Lesecoach mehr profitieren als von der Förderung durch die Eltern.

Von der Theorie in die Praxis: Implementation einer differenzierten Leseförderung in den schulischen Alltag unter dem Blickwinkel der Förderintensität

Natalie Förster, Henrike Mentel, Mareike Ehlert, Boris Forthmann, Elmar Souvignier
Westfälische Wilhelms-Universität Münster

Theoretischer Hintergrund

Die Lesekompetenz ist eine Schlüsselkompetenz für gesellschaftliche Teilhabe, deren wesentliche Grundlagen während der Grundschulzeit erworben werden. Zwar verfügen deutsche Viertklässler:innen im Mittel über durchschnittliche Lesekompetenzen, gleichzeitig sind die interindividuellen Unterschiede im internationalen Vergleich jedoch besonders groß. Für Lehrkräfte ergibt sich daher die Notwendigkeit, im Leseunterricht zu differenzieren.

Das Leseförderprogramm „Der Lese-Sportler“ wurde speziell zu diesem Zweck entwickelt. Das Programm integriert drei evidenzbasierte Fördermethoden zum genauen, flüssigen und verstehenden Lesen in einem Rahmenprogramm und bietet für jede Fördermethode unterschiedlich schwieriges Material an (Hebbecker et al., 2020).

Bislang wurde das Förderprogramm erfolgreich im Rahmen einzelner Forschungsprojekte im Regelunterricht eingesetzt (z. B. Peters et al., 2021). Der Einsatz neuartiger Unterrichtsmethoden erfordert jedoch auch immer eine Einarbeitung der Lehrkräfte. Angesichts der aktuellen Belastungssituation von Lehrkräften, fehlt diesen jedoch häufig die Zeit für eine Teilnahme an Fort- und Weiterbildungsprogrammen.

Das vorliegende Forschungsprojekt hat sich daher zum Ziel gesetzt, das Programm über Lehramtsstudierende im Praxissemester in die schulische Praxis zu implementieren. Dazu lernen Lehramtsstudierende vor ihrem Praxissemester in zwei universitären Lehrveranstaltungen das Lese-Sportler-Programm kennen. Während ihres Praxissemesters sollen sie Lehrkräfte dann bei der differenzierten Leseförderung von Kindern unterstützen. Ob und in welchem Ausmaß Kinder von diesem Förderangebot profitieren, wird dabei jedoch vor allem von der Förderintensität abhängen (Roberts et al., 2021).

Fragestellung

In unserem Beitrag beschäftigen wir uns mit der Implementation dieses Angebots in die schulische Praxis unter dem Blickwinkel der Förderintensität. Dazu quantifizieren wir einerseits das Ausmaß (zusätzlicher) differenzierter Leseförderung in Form der Häufigkeit und Dauer (Warren et al., 2007) sowie der Anzahl gelesener Wörter (Hammerschmidt-Snidarich et al., 2019), das ausgewählte Kinder während des Praxissemesters erhalten haben. Zudem analysieren wir, inwiefern die Förderintensität Varianz in den Lernzuwächsen der Kinder erklären kann. Wir erwarteten, dass die Häufigkeit, die Dauer und die Lesemenge die Lernzuwächse prädizieren.

Methode

Seit dem Sommersemester 2021 haben insgesamt 47 Grundschulstudierende in drei Praxissemesterdurchgängen an dem Projekt teilgenommen. Während des Praxissemesters haben sie ausgewählte Kinder in Kleingruppen mit dem Lese-Sportler-Material mindestens einmal pro Woche trainiert. Die einzelnen Fördereinheiten wurden von den Studierenden u.a. im Hinblick auf die Häufigkeit, Dauer und Anzahl gelesener Wörter sowie die eingesetzten Fördermethoden dokumentiert. Die Leistungsentwicklung der trainierten Kinder sowie deren Klassen wurde mit dem standardisierten Lesetest ELFE II (Lenhard et al., 2017) zum Halbjahresbeginn und -ende erhoben.

Die Auswertungen der Förderhäufigkeit und -dauer sowie der Lesemenge erfolgen deskriptiv. Anhand von hierarchischen multiplen Regressionsmodellen wird geprüft, inwiefern die Förderhäufigkeit, -dauer und die Lesemenge die Lernzuwächse vorhersagen können.

Ergebnisse

Da die Datenerhebungen noch laufen, liegen bislang nur vorläufige Ergebnisse vor. Von den 21 Studierenden aus dem ersten Praxissemesterdurchgang wurden insgesamt 168 Kinder aus 27 Klassen gefördert. Die Förderung verteilt sich dabei recht gleichmäßig auf Kinder aus den Klassenstufen 2, 3 und 4. Es wurden tendenziell eher leseschwache Kinder (T-Wert: M = 44,06; SD = 8,77) für die Förderung ausgewählt. Im Mittel wurden im Zeitraum von etwa Mitte Oktober 2021 bis Mitte Januar 2022 7,75 Fördereinheiten à 23,11 Minuten durchgeführt. Das entspricht einer

durchschnittlichen Förderung von insgesamt etwa 3 Stunden mit einem Maximalwert von 5,42 Stunden.

Diskussion

Die ersten Befunde deuten darauf hin, dass das Angebot zusätzlicher Leseförderung nur mit einer geringen Dosierung in den schulischen Alltag integriert wurde. Vergleicht man die bislang beobachtete Förderintensität, so liegt selbst die maximale Förderdauer mit 5,42 Stunden nur minimal über der kürzesten Interventionsdauer (5,33 Stunden) einer aktuellen Meta-Analyse zur Rolle der Förderintensität bei Leseförderprogrammen (Roberts et al., 2021). Internationale Vergleichsstudien zeigen, dass in Deutschland sowohl absolut als auch relativ vergleichsweise wenig Zeit für expliziten Leseunterricht aufgewandt wird. Im Beitrag diskutieren wir unsere Befunde im Hinblick auf den Stellenwert expliziten Leseunterrichts in Deutschland sowie die Implementation evidenzbasierter Fördermethoden in die schulische Praxis.

G-18: Die Professionelle Wahrnehmung von Lehrkräften fachspezifisch messen und fördern: Befunde aus vier Fächern

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* T03 R04 D10

Die Professionelle Wahrnehmung von Lehrkräften fachspezifisch messen und fördern: Befunde aus vier Fächern

Chair(s): **Rebekka Stahnke** (IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik), **Andreas Gegenfurtner** (Universität Augsburg)
*Diskutant*in(nen):* **Manfred Holodynski** (Westfälische Wilhelms-Universität Münster)

Die Professionelle Wahrnehmung stellt eine zentrale situationsspezifische Fertigkeit von Lehrkräften dar, die als Mediator zwischen ihrem Wissen und Überzeugungen und ihrem unterrichtspraktischen Handeln wirkt (Blömeke et al., 2022). Die Professionelle Wahrnehmung umfasst das *Noticing*, d.h. das Bemerken relevanter Aspekte einer Unterrichtsszene, sowie das *Knowledge-based Reasoning*, d.h. die wissensbasierte Deutung des Wahrgenommenen (Seidel & Stürmer, 2014). Während in den letzten Jahren besonders die Professionelle Wahrnehmung von generischen Unterrichtsfacetten untersucht wurde, rückt zunehmend die Wahrnehmung von fachspezifischen Aspekten in den Fokus. Die verschiedenen Kontexte und Inhalte, die mit konkreten Fächern verknüpft sind, stellen spezifische Anforderungen an Lehrkräfte. Inwiefern diese spezifischen Anforderungen bei der Entwicklung von Instrumenten zur Erhebung der fachspezifischen professionellen Wahrnehmung berücksichtigt werden müssen und inwiefern fachspezifische Ansätze zur Förderung der Professionellen Wahrnehmung notwendig sind, soll Gegenstand des Symposiums sein, insbesondere mit Blick auf die Didaktik des Sachunterrichts, der Mathematik, des Sportunterrichts und Deutsch als Fremd- bzw. Zweitsprache (DaF/DaZ). Während sich die ersten beiden Beträge methodischen Aspekten der Erhebung der Professionellen Wahrnehmung widmen, untersuchen Beitrag drei und vier die Wirksamkeit von Maßnahmen zur Förderung der Professionellen Wahrnehmung von Lehrkräften.

Der erste Beitrag vergleicht offene und geschlossene Erhebungsverfahren der Professionellen Wahrnehmung im Sachunterricht. Zusätzlich wird untersucht, inwiefern Gewissenhaftigkeit relevant für die Bearbeitung ist sowie welche kognitiven Prozesse bei der Bearbeitung ablaufen. Es zeigt sich u. a. nur ein geringer Zusammenhang zwischen beiden Erhebungsformaten sowie eine Bedeutsamkeit von Persönlichkeitsfacetten für die Bearbeitung des offenen Formats.

Der zweite Beitrag widmet sich in zwei Teilstudien der Frage, welche Rolle der mathematische als auch kulturelle Kontext für die vignettenbasierte Erfassung des Noticing spielt. Die Ergebnisse deuten darauf hin, dass das Noticing inhaltsübergreifend erfassbar ist, jedoch Kontexteffekte auf Ebene der Vignetten beobachtbar sind. Weiterhin sind für die Entwicklung kulturübergreifender Instrumente zusätzliche Validierungsschritte relevant.

Im dritten Beitrag wird die Wirksamkeit eines Praxissemesters mit begleitendem Seminar zur Förderung der situationsspezifischen Fähigkeiten für den inklusiven Sportunterricht untersucht. Die Intervention erwies sich als wirksam hinsichtlich der Förderung in Bezug auf die zentralen Anspruchsdimensionen eines inklusiven Sportunterrichts Anerkennung, individuelle Förderung und Teilhabe.

Der letzte Beitrag überprüft, wie sich die professionelle Wahrnehmung von sprachsensiblem Unterricht durch eine DaF/DaZ-didaktische Seminarteilnahme verändert. Mixed-Methods-Analysen

zeigten, dass viele Studierende sich hinsichtlich des Wahrnehmens sprachsensibler Unterrichtsaspekte verbessern. Zudem werden verschiede Aspekte der Sprachbildung unterschiedlich häufig adressiert.

Die Beiträge zeigen auf, dass fachspezifische Anforderungen sowie Kontextaspekte bei der Entwicklung von Instrumenten zur Erfassung der fachspezifischen Professionellen Wahrnehmung sowie bei deren Förderung berücksichtigt werden sollten.

Beiträge des Symposiums

Vergleich offener und geschlossener Erhebungsverfahren der professionellen Unterrichtswahrnehmung von Lernunterstützung im naturwissenschaftlichen Sachunterricht

Isabel Eßling[1], Cornelia Sunder[1], Lena Steidtmann[2], Mirjam Steffensky[2], Nicola Meschede[1]
[1]Westfälische Wilhelms-Universität Münster, [2]Universität Hamburg

Theoretischer Hintergrund

Professionelle Unterrichtswahrnehmung gilt als wichtige situationsspezifische Kompetenz von Lehrkräften, da sie als Mediator zwischen dem Professionswissen und der Performanz im Unterricht angenommen wird (Blömeke et al., 2022; Seidel & Stürmer, 2014). Bisherigen Studien zur professionellen Wahrnehmung liegt häufig eine kontextualisierte Erfassung anhand authentischer Unterrichtssituationen, z.B. mithilfe von Videoszenen, zugrunde (Barth, 2017; Seidel & Thiel, 2017). Dabei unterscheiden sich die eingesetzten Instrumente jedoch in der Erfassung der professionellen Wahrnehmung; so werden geschlossene Ratingitems, offene Aufgabenstellungen oder eine Kombination aus beidem verwendet. Erste Befunde zur Erfassung professioneller Wahrnehmung von Klassenführung deuten darauf hin, dass offene und geschlossene Erhebungsformate nur gering bis gar nicht korrelieren (Gold et al., 2016; König et al., 2014). Bislang ist unklar, ob dieser Befund über verschiedene Domänen und Unterrichtsqualitätsmerkmale hinweg generalisierbar ist. Zudem gibt es kaum Erkenntnisse dazu, wodurch sich mögliche Unterschiede ergeben. Während sich bei der Präsentation geschlossener Ratingitems beispielsweise eine starke Aufmerksamkeitslenkung nachteilig auf die genaue Erfassung der Wahrnehmung auswirken könnte, werden durch den niedrigeren Grad der Standardisierung einer offenen Erfassung möglicherweise andere Merkmale wie eine gewissenhafte Arbeitsweise miterfasst. Ergänzende qualitative Analysen, die Einblick in kognitive Prozesse während der Bearbeitung geben, könnten Potential für die Klärung von (fehlenden) Zusammenhängen bieten.

Fragestellungen

Im Rahmen der vorliegenden Studie werden deshalb die offene und geschlossene Erfassung professioneller Unterrichtwahrnehmung von Lernunterstützung im Kontext des naturwissenschaftlichen Sachunterrichts genauer untersucht. Dabei stehen folgende Forschungsfragen im Fokus:

F1. Wie hängen eine offenes und ein geschlossenes videobasiertes Instrument zur Erfassung der professionellen Wahrnehmung von Lernunterstützung im naturwissenschaftlichen Grundschulunterricht zusammen?

F2. Hängt eine gewissenhafte Arbeitsweise mit der Messung der professionellen Wahrnehmung zusammen?

F3. Welche kognitiven Prozesse laufen bei der Bearbeitung der videobasierten Instrumente zur Erfassung der professionellen Wahrnehmung ab?

Methode

Zur Beantwortung von F1 analysierten N=110 Lehramtsstudierende (Semester M= 5.05, SD= 0.79) eine identische Unterrichtszene aus dem Sachunterricht zum Thema Aggregatzustände sowohl offen, in Form der Beschreibung und Interpretation relevanter lernunterstützender Ereignisse, als auch geschlossen, d.h. mithilfe von likert-skalierten Ratingitems (trifft nicht zu, trifft eher nicht zu, trifft eher zu, trifft zu). Das Unterrichtsvideo und die verwendeten Ratingitems entstammen einem bereits validierten Videoinstrument (Meschede et al., 2015). Die standardisierten Ratingitems wurden anhand eines vorliegenden expertenvalidierten Masterratings bewertet. Je Item wurden 2 Punkte für die exakte Übereinstimmung und 1 Punkt für tendenzielle Übereinstimmung mit dem Masterrating vergeben. Die offene Analyse wurde qualitativ inhaltsanalytisch ausgewertet und durch zwei Rater (35% Doppelkodierungen, κ= 0.89) kodiert. Das Erkennen einer lernunterstützenden Maßnahme wurde in Übereinstimmung mit einem Masterrating mit 1 Punkt bewertet und das Interpretationsniveau je nach Elaboriertheit mit 0-2 Punkten. Die Gewissenhaftigkeit (F2) wurde mittels der Subskala des NEO-Fünf-Faktoren-Inventars (NEO-FFI; Borkenau & Ostendorf, 2008) erfasst. Zusätzlich wurden leitfadengestützte, teil-standardisierte Laut-Denken-Interviews während der Bearbeitung der offenen und geschlossenen Analyse mit N=16 Lehramtsstudierenden durchgeführt (F3), die derzeit inhaltsanalytisch ausgewertet werden.

Ergebnisse

Die Ergebnisse zeigen eine niedrige Korrelation der beiden Erhebungsformate (r=.20, p=.050*). Im Mittel erreichten die Studierenden bei der offenen Analyse Moffen= 59% der maximal erreichbaren Punkte (SDoffen= 18%) und im geschlossenen Format Mgeschlossen= 49% (SDgeschlossen= 18%). Die Gewissenhaftigkeit der angehenden Lehrkräfte korrelierte schwach mit ihrer Leistung in der offenen Videoanalyse (r=.21, p=.036*). Zusammenhänge zur Bearbeitung der geschlossenen Rating-Items konnten hingegen nicht gezeigt werden (r=.14, p=.181).

Diskussion

Die Befunde bestätigen Ergebnisse aus anderen Domänen und deuten somit auf eine Generalisierbarkeit niedriger Zusammenhänge zwischen offen- und geschlossen-erfasster professioneller Wahrnehmung hin. Dies unterstreicht die Notwendigkeit für die Ergründung von Ursachen. Diesbezüglich zeigt die vorliegende Studie, dass Unterschiede in der gewissenhaften Arbeitsweise bei der offenen Messung der professionellen Wahrnehmung eine Rolle spielen, während die Leistungsvarianz in der geschlossenen Videoanalyse unabhängig von dieser Persönlichkeitsfacette ist. Die derzeitige Auswertung der Laut-Denken-Interviews kann weitere Erkenntnisse zu möglichen Ursachen liefern, die auf der Tagung präsentiert werden.

Welche Rolle spielen Kontexte bei der Erfassung von Teacher Noticing zu Qualitätsmerkmalen von Mathematikunterricht?

Marita Friesen[1], **Anika Dreher**[2]
[1]Pägagogische Hochschule Heidelberg, [2]Pädagogische Hochschule Freiburg

Theoretischer Hintergrund

In komplexen Unterrichtssituationen lernrelevante Ereignisse identifizieren und auf Grundlage professionellen Wissens interpretieren zu können, gilt als zentraler Aspekt von Lehrkräfteexpertise (Stahnke et al., 2016; van Es & Sherin, 2002). Teacher Noticing adressiert häufig wichtige Qualitätsmerkmale von Mathematikunterricht, wenn Lehrkräfte in Unterrichtssituationen beispielsweise das mathematische Denken von Lernenden oder schwierige Darstellungswechsel identifizieren und interpretieren sollen, um entsprechend adaptiv handeln zu können (Dreher & Kuntze, 2015). Mit einer zunehmenden Studienanzahl in vielfältigen mathematischen und kulturellen Kontexten gewinnt die Frage nach deren Rolle für die Erfassung von Noticing an zentraler Bedeutung. Die Rolle mathematischer Kontexte für Noticing wurde bereits früh adressiert mit der Frage, ob Noticing aus dem Bruchrechnen auf den Bereich Algebra übertragbar sei (Sherin et al., 2011). Obwohl diese Frage für die Lehrkräftebildung zentral scheint (Dindyal, 2021), gab es im Folgenden kaum Studien zum Vergleich von Noticing zwischen verschiedenen mathematischen Inhaltsbereichen. Die Frage nach der Bedeutung von kulturellen Kontexten stellt sich bei der Entwicklung (vignettenbasierter) Testinstrumente für internationale Vergleichsstudien (vgl. Dindyal et al., 2021). Kulturelle Einflüsse können eine wesentliche Bedeutung dafür haben, was als qualitätsvoller Mathematikunterricht angesehen wird (vgl. Louie, 2018) und damit auch, was in einer Unterrichtssituation als lernrelevant identifiziert und interpretiert wird. Allerdings gibt es noch wenige Studien, die systematisch die Entwicklung interkulturell valider (vignettenbasierter) Testinstrumente im Bereich Noticing adressieren.

Fragestellung

Die skizzierten Forschungsdesiderate werden in zwei Studien unter folgenden ausgewählten Forschungsfragen aufgegriffen:

Studie 1 (Friesen & Kuntze, 2021): Welche Rolle spielt der mathematische Kontext für das Noticing von Mathematiklehrkräften zum Qualitätsmerkmal „Umgang mit vielfältigen Darstellungen“ in den Inhaltsbereichen Brüche und Funktionen?

Studie 2 (Dreher et al., 2021): Welche Rolle spielt der kulturelle Kontext für das Noticing von Fachdidaktik-Expert:innen aus Taiwan und Deutschland zum Qualitätsmerkmal „Umgang mit mathematischem Denken von Lernenden“?

Methode

In Studie 1 nahmen N=175 Mathematiklehrkräfte teil, in Studie 2 N=38 Professor:innen aus der Fachdidaktik Mathematik (jeweils 19 aus Taiwan und Deutschland). In beiden Studien wurden vignettenbasierte Testinstrumente entwickelt, um das Noticing zum jeweiligen Qualitätsmerkmal von Mathematikunterricht zu erfassen. Hierfür enthalten die Testinstrumente 12 Vignetten (Studie 1: jeweils 6 Vignetten zu Brüchen und Funktionen) bzw. 18 Vignetten (Studie 2). Die Vignetten repräsentieren jeweils Unterrichtssituationen (z.B. Textformat), in denen eine Lehrer-Schüler-

Interaktion zum jeweiligen Qualitätsmerkmal gezeigt wird. Die Reaktion der Lehrkraft enthält dabei jeweils mindestens einen fachdidaktischen Normbruch, den es zu identifizieren und zu interpretieren gilt (vgl. Herbst & Kosko, 2014). In Studie 2 erfolgte die Entwicklung der Vignetten iterativ (4 Phasen) und kollaborativ zwischen dem taiwanesischen und dem deutschen Projektteam, um die interkulturelle Validität der Vignetten zu gewährleisten (vgl. Dreher et al., 2021). Die offenen Antworten wurden in beiden Studien nach den jeweils adressierten Qualitätsmerkmalen kodiert (Studie 1: $\kappa \geq 0.89$; Studie 2: $\kappa \geq 0.72$). In Studie 1 konnten die erhaltenen Scores raschskaliert werden, wobei eine Gruppierung auf die beiden Subdimensionen Brüche bzw. Funktionen zum besten Modellfit führte ($0.96 \leq wMNSQ \leq 1.04$; $-0.9 \leq T \leq 1.2$; vgl. Brandt, 2008).

Ergebnisse und Diskussion

Die Studien zeigen wichtige Erkenntnisse zur Rolle von Kontexten bei der vignettenbasierten Erfassung von Noticing: In Studie 1 ergaben sich zwischen den beiden Inhaltsbereichen (Brüche, Funktionen) keine signifikanten Unterschiede in den empirischen Itemschwierigkeiten. Noticing zu übergreifenden Qualitätsmerkmalen (hier: Umgang mit Darstellungen) könnte damit inhaltsübergreifend erfassbar und ggf. auch entwickelbar sein. Allerdings zeigten die Vignetten signifikante Schwierigkeitsunterschiede innerhalb der beiden Inhaltsbereiche, was zusätzlich auf die Bedeutung des Kontexts auf Vignettenebene hinweist. In Studie 2 führte die mehrphasige Validierung der Vignetten dazu, dass die meisten taiwanesischen und deutschen Expert:innen den konstruierten Normbruch identifizieren konnten, wobei bei dessen Interpretation jedoch kulturelle Sichtweisen (z.B. Prozess vs. Prozedur) sichtbar wurden. Die Autoren schlagen die Entwicklung eines kulturspezifischen Kodiersystems für internationale Vergleichsstudien im Bereich Teacher Noticing vor.

Das „missing Link“ zwischen Dispositionen und Performanz: Weiterentwicklung situationsspezifischer Fähigkeiten für den inklusiven Sportunterricht

Jan Erhorn[1], **Wiebke Langer**[2]
[1]Universität Osnabrück, [2]Universität Hamburg

Theoretischer Hintergrund

In der Lehrkräftebildung wurde die zentrale Bedeutung situationsspezifischer Fähigkeiten für ein professionelles Handeln erkannt. So bilden die Prozesse des Wahrnehmens, Interpretierens und Entscheidens eine wichtige Brücke für die Verknüpfung von anforderungsrelevanter Dispositionen mit der unterrichtlichen Performanz (Blömeke, Gustafsson & Shavelson, 2015). Für die Förderung situationsspezifischer Fähigkeiten wurde die besondere Bedeutung von videografischer Fallarbeit nachgewiesen (Gold, Förster & Holodynski, 2013).

Da sich der Sportunterricht, u. a. durch spezifische Fachräume und Inhalte, die Fokussierung auf Bewegung sowie damit verknüpfte motorische und soziale Prozesse, von anderen Schulfächern unterscheidet, ergeben sich fachspezifische Anforderungen an die situationsspezifischen Fähigkeiten von Sportlehrkräften, die bei der Entwicklung von Lehrformaten berücksichtigt werden müssen (Erhorn, Langer & Möller, 2020). Allerdings steht die Entwicklung von Lehrformaten, welche auf die Vermittlung situationsspezifischer Fähigkeiten gerichtet sind, in der Sportdidaktik erst am Anfang.

Die im Rahmen des vom BMBF geförderten Projekts „Qualifizierung angehender Sportlehrkräfte für einen inklusiven Sportunterricht (QiS)“ entwickelte, aus einem Seminar und einem begleiteten Praxissemester bestehende Lehrsequenz, setzt auf die Bearbeitung typischer Anforderungssituationen eines inklusiven Sportunterrichts in Form von Fall- und Portfolioarbeit. Dabei werden anforderungsbezogen situationsspezifische Fähigkeiten gefördert und relevante Dispositionen vermittelt (Erhorn & Langer, 2022).

Fragestellung

Der Vortrag fokussiert die Fragestellung, inwieweit durch Video-Fallarbeit im Rahmen eines kompetenzbasierten, an typischen Anforderungssituationen orientierten Seminars situationsspezifische Fähigkeiten für einen inklusiven Sportunterricht gefördert werden können.

Methode

Die Wirksamkeit des Seminars wurde in Form eines Prä-Post- und Interventions-Kontrollgruppen-Designs (IG N=114; KG N=142) überprüft. Dafür wurden die selbstentwickelten Testskalen zur Erfassung situationsspezifischer Fähigkeiten Wahrnehmen und Interpretieren mit dem Fokus auf die inklusiven Anspruchsdimensionen Anerkennung (ViProQiS_A), individuelle Förderung (ViProQiS_IF) sowie Teilhabe (ViProQiS_T) (Langer, Bruns & Erhorn, 2022, eingereicht a, eingereicht b) an zwei Messzeitpunkten, vor und unmittelbar nach dem Seminar, eingesetzt. Die Auswertung erfolgte mit Hilfe varianzanalytischer Verfahren.

Ergebnisse und ihre Bedeutung

Das entwickelte Seminar erwies sich als wirksam. So wiesen Studierende der Interventionsgruppe gegenüber Studierenden der Kontrollgruppe bedeutsame Zuwächse in den situationsspezifischen Fähigkeiten des Wahrnehmens und Interpretierens in den Kompetenzbereichen Anerkennung ($F = 117.902$, $df = 1$, $p < .001$), Teilhabe ($F = 23.293$, $df = 1$, $p < .001$) und individuelle Förderung ($F = 43.053$, $df = 1$, $p < .001$) auf.

Die Ergebnisse geben eindeutige Hinweise, darauf dass mit der Videofallarbeit in der entwickelten Unterrichtssequenz die situationsspezifischen Fähigkeiten Wahrnehmen und Interpretieren von angehenden Lehrkräften in Bezug auf die zentralen Anspruchsdimensionen eines inklusiven Sportunterrichts Anerkennung, individuelle Förderung und Teilhabe sehr wirksam gefördert werden können. Die entwickelte Lehrsequenz wurde inzwischen bereits an mehreren universitären Standorten implementiert, sodass Auswirkungen auf die Sportlehrkräftebildung zu erwarten sind.

Fachspezifisch professionell wahrnehmen – Deutsch als Zweitsprache und sprachsensibler Unterricht im Fokus

Christine Stahl, Kristina Peuschel
Universität Augsburg

Professional Vision etabliert sich zunehmend auch mit fachspezifischen Fragestellungen als Forschungs- und Lehrgegenstand in Lehrkräftebildung und empirischer Bildungsforschung. Im Fach Deutsch als Zweitsprache (DaZ), das den Deutschunterricht für neuzugewanderte Schüler:innen zum Gegenstand hat, sowie im sprachsensiblen Fachunterricht, der auf die sprachlich-fachliche Bildung aller Schüler:innen fokussiert, sind die sprachlernförderlichen, interaktionalen Handlungskompetenzen

von Lehrkräften besonders relevant. Diese im Lehramtsstudium mittels professioneller, fachspezifischer Wahrnehmung von Unterricht in sprachlich heterogenen Klassen anzubahnen, ist Gegenstand der fachdidaktischen Ausbildung im Fach DaZ an der Universität Augsburg.

In der vorgestellten Studie zur professionellen Unterrichtswahrnehmung von Aspekten sprachsensiblen Unterrichts wird der Frage nachgegangen, wie angehende Lehrkräfte sprachsensiblen, videographierten Unterricht professionell wahrnehmen. Weitere Analysefragen beziehen sich konkret darauf, wie angehende Lehrkräfte wahrgenommene Aspekte wahrnehmen, beschreiben, erklären und vorhersagen. Die Stichprobe umfasst n = 36 angehende Grund- und Mittelschullehrkräfte, die mit einer Ausnahme DaZ als Unterrichtsfach studieren. Im Durchschnitt befinden sich die Studierenden im fünften Semester und besuchten ca. drei Seminare im Fach DaZ vorab. Die Datenerhebung fand im Rahmen des Seminars „Unterrichtsinteraktion unter der Perspektive von Sprachförderung und sprachlich-fachlicher Bildung“ in zwei Kohorten im Sommersemester 2020 sowie im Wintersemester 2020/21 statt. Als Erhebungsinstrument wurde der qualitative, onlinebasierte Test ‚sHiVa – sprachliche Heterogenität in Videos analysieren‘ (vgl. Stahl 2022) eingesetzt. Jeweils in der ersten und letzten Seminarsitzung bearbeiteten die Studierenden den Test, welcher alle Teilprozesse der professionellen Wahrnehmung – wahrnehmen, beschreiben, erklären und vorhersagen – widerspiegelt. Sechs vorgegebene Beobachtungsschwerpunkte, welche als strukturierte Beobachtungshilfe definiert werden können, unterstützen die angehenden Lehrkräfte, ihr Augenmerk auf sprachsensible Aspekte zu legen. Die Studierenden analysieren im Rahmen des Tests zwei ca. zehnminütige videographierte Unterrichtsmitschnitte aus dem ViKo_spracHe (vgl. Stahl & Peuschel, i.V.). Die Unterrichtsmitschnitte „Hundertwasser“ und „Weltall“ zeigen authentischen, sprachsensiblen Fachunterricht aus der Primarstufe und der Sekundarstufe I.

Die Auswertungen der Videoanalysen der angehenden Lehrkräfte erfolgten in einem mixed-methods-Design. Dabei wurden die Wahrnehmungen quantitativ ausgewertet. Die Berechnung der Häufigkeiten und t-Tests bei abhängigen Stichproben zeigten individuelle und vom Beobachtungsschwerpunkt abhängige Unterschiede in der Fähigkeit, sprachsensiblen Unterricht in sprachlich heterogenen Klassen professionell wahrzunehmen. Über zwei Drittel der Stichprobe nehmen im zweiten gegenüber dem ersten Messzeitpunkt mehr sprachsensible Aspekte wahr. Bei ca. einem Drittel verändert sich die Anzahl der wahrgenommenen Items nicht (n = 4) bzw. nimmt ab (n = 11).

Qualitative inhaltsanalytische Auswertungen im Bereich des knowledge based reasonings zeigen, dass die Proband:innen beim Beschreiben, Erklären und Vorhersagen ihren Beobachtungsfokus unterschiedlich legen. Die Mehrheit fokussiert die Handlungen der Lehrkraft oder der Schüler:innen. Nicht so häufig werden Interaktionen zwischen der Lehrkraft und einzelnen Schüler:innen oder Peer-Interaktionen thematisiert. Im Rahmen des Erklärens gelingt es den Studierenden weniger einen Bezug zu Theorien oder empirischen Studien herzustellen. Wenige referieren Studienergebnisse, nennen Modelle oder thematisieren theoretische Inhalte aus dem Seminar. Wirft man einen Blick auf die angesprochenen DaZ-spezifischen Themen, so wird deutlich, dass viele Proband:innen in ihren Analysen Wortschatzarbeit, die Sprachkompetenz der videographierten Schüler:innen, sprachlernförderliche Rückmeldungen, Visualisierungen oder die Lehrsprache ansprechen. Seltener

wird auf die Fertigkeit Lesen, die Vermittlung von Grammatik oder die Ausspracheschulung eingegangen. Individuelle Unterschiede seitens der Proband:innen zeigen sich daneben in der Verwendung unterrichtsspezifischer Sprache.

Ziel der Studie ist es, die fachspezifische professionelle Unterrichtswahrnehmung angehender Lehrkräfte im Fach Deutsch als Zweitsprache zu untersuchen und weiterzuentwickeln. Dabei soll DaZ-spezifisches anwendungsorientiertes, theoretisches Wissen über qualitätsvollen sprachlernwirksamen Unterricht mit dem späteren Handlungsfeld in sprachlich heterogenen Schulklassen verknüpft werden. Die Analyseergebnisse, insbesondere des noticings, zeigen, welche Aspekte eines sprachsensiblen Fachunterrichts seitens der Studierenden besonders gut bzw. weniger gut wahrgenommen werden. Erste Auswertungen im Bereich des knowledge based reasonings zeigen zum einen, auf welche am Unterricht beteiligten Personen, zum anderen auf welche DaZ-spezifischen Themen die angehenden Lehrkräfte ihren Beobachtungsfokus legen. Die Ergebnisse geben unter anderem Hinweise darauf, welche Themen in der DaZ-Lehrkräftebildung stärker thematisiert werden können.

G-19: „Bedrohung“ der Testwertinterpretation durch Testmotivation? – Fragen der Validität in der Kompetenzforschung

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* R11 T00 D03

„Bedrohung“ der Testwertinterpretation durch Testmotivation? – Fragen der Validität in der Kompetenzforschung

Chair(s): **Christina Watson** (SRH in NRW, Deutschland), **Daniel Scholl** (Universität Siegen)
*Diskutant*in(nen):* **Niclas Schaper** (Universität Paderborn)

In den vergangenen drei Jahrzehnten hat die Professionsforschung die Modellierung von latenten Kompetenzkonstrukten und deren Erfassung in handlungsnahen Situationen zusammen mit der begleitenden Auseinandersetzung mit Grundsatzfragen der Validität (Hartig & Frey, 2012; Jenßen et al., 2015) im neueren Sinne der Testwertinterpretation (Kane, 2013) intensiviert. Speziell bei der Operationalisierung von komplexen Konstrukten wurde dabei zunehmend die Herausforderung gesehen, bei der Prüfung der unterschiedlichen Validitätsaspekte (Messick, 1995) auch testökonomische Entscheidungen zum Beispiel zu Itemformaten zu berücksichtigen, deren Effekte auf die Qualität der Testergebnisse vor allem bei zeit- und arbeitsaufwendigen Low-Stakes Assessments groß sein können (Baumert & Demmrich, 2001): So werfen beispielsweise fehlende Konsequenzen für Testteilnehmende und hohe Anforderungen an deren freiwilliges Engagement die Frage auf, inwiefern Tests weniger das eigentliche Konstrukt als vielmehr die Testmotivation erfassen (Sanchez et al., 2000).

In diesem Symposium wird diese Validierungsfrage bezogen auf forschungspragmatische Itemformatentscheidungen aufgegriffen: Mithilfe von ausgewählten Kompetenzstudien mit offenen und geschlossenen Items sollen weiterführende Testbearbeitungsindikatoren zur praktikablen Erfassung der Testmotivation (z.B. Green, 2008; Eklöf, 2008) und erweiterten Validitätsprüfung gesucht werden.

In den Beiträgen 1 bis 3 werden zunächst eng verwandte, komplexe Konstrukte der professionellen Lehrinnen- und Lehrerkompetenz und deren (vignettenbasierte) Operationalisierung über unterschiedliche Itemformate samt Validierungsbefunden vorgestellt: In Beitrag 1 wird ein geschlossener Onlinetest zur Erfassung der situationsspezifischen Fähigkeit zum Treffen wechselwirkender Planungsentscheidungen berichtet. Obwohl dieser Test erfolgreich faktoriell validiert wurde, machen festgestellte signifikante Zusammenhänge zwischen der Bearbeitungszeit und der Testleistung eine eingehendere Betrachtung etwaiger Validitätseinschränkungen durch die Testmotivation notwendig.

In Beitrag 2 wird ein offener Onlinetest zur Erfassung der Reflexionsperformanz von Lehramtsstudierenden und seine gründliche Prüfung auf inhaltliche, konvergente und divergente Validität präsentiert. Anzeichen auf mögliche negative Effekte von fiktiven Textvignetten lassen allerdings die Berücksichtigung zusätzlicher Motivationsmaße notwendig erscheinen, um die Testwerte zur Reflexionskompetenz fundierter unter dem Validitätsaspekt interpretieren zu können.

In Beitrag 3 wird die Implementierung einer geschlossenen Online-Version eines Paper-Pencil-Tests mit gemischten Aufgabenformaten zum bildungswissenschaftlichen Wissen von Lehramtsstudierenden besprochen. Auch wenn die IRT-Skalierung zu einem Implementierungserfolg des ursprünglich bereits curricular und kriterial validen Test geführt hat, verlangen die signifikant

geringeren Itemschwierigkeiten des geschlossenen Tests nach einer eingehenderen Validitätsprüfung etwaiger aufgabenbedingter Testmotivationseffekte.

In Beitrag 4 werden diese Tests mit einem Schwerpunkt auf den Aufgabenformaten bei ihrem kombinierten Einsatz in einer gemeinsamen Studie schließlich auf aussagekräftige Testmotivationsindikatoren zur erweiterten Validitätsprüfung hin analysiert. In der Diskussion werden schließlich Chancen und Grenzen der Erweiterung von Verfahren zur Validitätsprüfung durch Maße der Testmotivation abgewogen und Implikationen für die Testentwicklungspraxis angedacht.

Beiträge des Symposiums

Erfassung der Fähigkeit zum interdependenten Planungsentscheiden mit einem standardisierten Vignettentest

Daniel Scholl[1], Simon Küth[1], Christoph Schüle[2]
[1]Universität Siegen, [2]NLQ

In der Planung von Unterricht, verstanden als „didaktische Reflexion“ (David, 2018, 74) und „‚anticipatory reflection' as future-oriented reflection before action“ (Conway, 2001, S. 90), wird eine wichtige Grundlage für die Weiterentwicklung seiner Qualität gesehen (Wengert, 1989; Yinger & Hendricks-Lee, 1995). Es wird allerdings angenommen, dass nur eine professionelle Unterrichtsplanung eine solche Grundlage bietet (König et al., 2015), sodass die Förderung der situationsspezifischen Planungsfähigkeit als zentrale Aufgabe an die Lehrerbildung gerichtet ist (KMK, 2019). Die Professionalität dieser Fähigkeit zeigt sich dabei unter anderem daran, dass (angehende) Lehrkräfte ihre zu treffenden Planungsentscheidungen am Planungsprinzip der Interdependenz ausrichten können, um gedanklich eine kohärente Handlungsstruktur ihres künftigen Unterrichts zu entwerfen. Dieses Prinzip, das als eine „zentrale generische Anforderung“ (König et al. 2015, S. 378) der Unterrichtsplanung gilt und zu ihren bedeutendsten Metakriterien zählt (Vogelsang & Riese, 2017), verlangt, über die Stimmigkeit aller Planungsentscheidungen nachzudenken und beispielsweise zu klären, ob die Entscheidungen für Lernziele und für Methoden kohärent zueinander und zu den individuellen Lernvoraussetzungen getroffen wurden. Die Bedeutung, die diesem Prinzip zugesprochen wird, zeigt sich auch daran, dass entsprechende Anforderungen an die wechselseitige Verknüpfung aller Planungsentscheidungen sowohl in didaktischen Ansätzen (Scholl, 2018), Studien des Decision-Making-Paradigmas (Shavelson & Stern, 1981) und gegenwärtigen fachübergreifenden (s. zur Übersicht Rothland, 2021) und fachbezogenen (z.B. Sachunterricht: Kirsch, 2020; Physik: Schröder et al., 2020) Kompetenzmodellen formuliert werden.

Im Gegensatz zu dieser zugeschriebenen Bedeutung fällt auf, dass bisher ein Rahmenmodell zur Spezifikation und ein Test zur Messung der kognitiven Operation des interdependenten Planungsentscheidens fehlen. Beides wäre aber notwendig, um das gut belegte Problem aufzuklären, dass viele angehende Lehrkräfte insbesondere mit der Herstellung von Kohärenz in ihren Unterrichtsplanungen Probleme zu haben scheinen (John, 2006; Koeppen, 1998; Mutton et al., 2011), und Konsequenzen für die Förderung interdependenten Planungsentscheidens abzuleiten.

Um diesem Desiderat zu begegnen, soll ein entsprechendes zweidimensionales Rahmenmodell mit einer formalen und einer inhaltlichen Dimension und dessen Operationalisierung vorgestellt werden: Dieses Modell schließt an die Annahme der wissensbasierten situationsspezifischen

Planungsfähigkeiten des Wahrnehmens, Interpretierens und Entscheidens von Planungsoptionen an (Blömeke et al., 2015; König et al., 2017). Die formale Dimension betrifft die wechselseitige Verknüpfung der ziel-, themen-, methoden- und medienbezogenen Wissenselemente beim situationsbezogenen Planungsentscheiden (Tillema, 1984; Zierer & Seel, 2012), die inhaltliche das Konzeptions- und Planungswissen als Basis dieser Fähigkeit (Baumert & Kunter, 2006).

Unter Rückgriff auf die Stärken von vignettengestützten Testverfahren (Friesen & Feige, 2020) und bei Beachtung planungstestkritischer Aspekte (Rothland, 2021), wurde dieses Modell im Rahmen einer rationalen Testkonstruktionsstrategie (Moosbrugger & Brandt, 2020) über einen standarisierten vignettenbasierten Online-Fähigkeitstest mit geschlossenen Items operationalisiert: Um interdependentes Entscheiden ökonomisch und analytisch scharf als Teilprozess der Unterrichtsplanung zu erfassen, umfasst dieser Test insgesamt zwölf Vignetten, für die jeweils diejenige Planungsoption auszuwählen ist, die eine initiale Unterrichtsplanungssituation am kohärentesten berücksichtigt (z.B. „Darstellung statistischer Daten“, Mathematik, Gymnasium Sek. I).

Zur Absicherung der Inhaltsvalidität wurde der Test in einem dreischrittigen Verfahren zusammen mit kriteriengeleitet gewählten Expert*innen entwickelt. Pilotiert wurde er in einer Gesamtstichprobe von N = 547 Personen, davon 126 angehende und 172 praktizierende Lehrkräfte. Die Prüfung der Konstruktvalidität (Hartig et al., 2020) ergab, dass die postulierten Fähigkeitsdimensionen empirisch in einem mehrdimensionalen Raschmodell (multidimensionales within-item-Modell mit bestem Fit und wMNSQ im Bereich von 0,91 ≤ wMNSQ ≤ 1,1) bei akzeptablen Reliabilitäten (EAP im Bereich von 0,57 ≤ EAP ≤ 0,67) abbildbar sind.

Trotz dieser ersten Validitätshinweise hat sich sowohl in der Pilotierungsstudie als auch beim Einsatz des Tests in einer unabhängigen Evaluationsstudie mit zwei Messzeitpunkten gezeigt, dass es signifikante Zusammenhänge zwischen der Bearbeitungszeit und der Testleistung gibt. Diese Zusammenhänge lassen die Frage aufkommen, ob die Anforderungen an die Testmotivation, die das lese- und konzentrationsintensive Durchdenken der komplexen Vignetten stellt, Einschränkungen für die Validität bedeuten könnten.

Erfassung der Reflexionskompetenz angehender Lehrkräfte mittels offener Situationsvignetten

Jana Meier, Christoph Vogelsang
Universität Paderborn

Im Kontext der Professionalisierung von (angehenden) Lehrkräften wird Reflexion neben anderen Begründungsfiguren (Klempin, 2019) als „Schlüsselkonzept für den Aufbau von Wissen und den Transfer von pädagogischem Wissen zum Handeln“ (Dittrich, 2020, S. 71) aufgefasst (z.B. Häcker, 2017; Korthagen, 2001; Leonhard & Rihm, 2011; Neuweg, 2005). Reflexion soll in ihrer Relationierungsfunktion zwischen Theorie und Praxis (Tulodziecki, Herzig & Blömeke, 2017) einen Raum bieten, den Umsetzungserfolg von Theorien in der Praxis abzugleichen (Leonhard & Abels, 2017), um so einen Beitrag zur Förderung und Weiterentwicklung professioneller Handlungskompetenz zu leisten. Legt man ein umfassendes Kompetenzverständnis zugrunde, so lässt sich professionelle Handlungskompetenz als Kontinuum zwischen den Dispositionen einer Person, ihren situationsspezifischen Fähigkeiten und dem beobachtbaren Verhalten (Performanz) definieren (Blömeke, Gustafsson & Shavelson, 2015). Überträgt man dieses Kompetenzverständnis

auf den Reflexionsbegriff (Aufschnaiter, Fraij & Kost, 2019), zeichnet sich Reflexionskompetenz ebenfalls durch kognitive und affektiv-motivationale Dispositionen (z.B. Wissen über und Einstellungen zu Reflexion) aus, die vermittelt über reflexionsbezogene Denkprozesse in Reflexionsperformanz (z.B. Verschriftlichungen) transformiert werden.

Die Operationalisierung und Messung von Reflexionskompetenz gestalten sich seit langem als zentrale Herausforderungen für die empirische Bildungsforschung. In Anlehnung an das Kontinuumsmodell von Kompetenz (Blömeke et al., 2015) und ein kognitiv-psychologisches Verständnis (Aeppli & Lötscher, 2016; Korthagen, 2001) lassen sich reflexionsbezogene Denkprozesse – als die situationsspezifischen Fähigkeiten von Reflexionskompetenz – durch konkrete Denkgegenstände (Inhalte der Reflexion: Handlungen, Erfahrungen, Beobachtungen…), unterschiedliche kognitive Aktivitäten sowie Ziele charakterisieren (z.B. Beauchamp, 2006; Hatton & Smith, 1995; Leonhard, 2013). Da die Denkprozesse einer Person jedoch nicht direkt erfassbar sind, ist es gängig, einen „Umweg“ über die Reflexionsperformanz zu gehen und Rückschlüsse auf die dahinterliegenden Prozesse zu ziehen.

In Anlehnung an die theoretischen Modelle und Perspektiven auf Reflexion, wurde ein offener Vignettentest zur Erfassung der Reflexionsperformanz Lehramtsstudierender entwickelt und validiert. Zum einen wurde der Test mittels einer Expert*innenbefragung inhaltsvalidiert (N=16, aus verschiedenen Disziplinen). Hierbei wurde die Realitätsnähe der Situaionsbeschreibungen sowie die Eindeutigkeit mittels einer fünfstufigen Likert-Skala bewertet und offenes Feedback für die Überarbeitung der Vignetten gegeben. Darüber hinaus wurden die Vignetten gemeinsam mit einem Test zur Erfassung des bildungswissenschaftlichen Professionswissens (Seifert & Schaper, 2010) sowie Skalen zur Erfassung reflexionsbezogener Selbstwirksamkeitserwartungen (Fraij, 2018) und Einstellungen zu Reflexion (Neuber & Göbel, 2016) im Rahmen einer Onlinebefragung eingesetzt (N = 295). Durch Korrelationsanalysen wurden erwartungskonforme Zusammenhänge zwischen dem Professionswissen, der subjektiven Einschätzung der Relevanz von Reflexion im Studium, der reflexionsbezogenen Selbstwirksamkeitserwartungen und der Reflexionsperformanz nachgewiesen. Unterschiede in der Reflexionsperformanz können statistisch signifikant durch das Professionswissen der Studierenden erklärt werden. Die erhobenen affektiv-motivationalen Dispositionen können hingegen keine Unterschiede in der Reflexionsperformanz der Studierenden erklären, dennoch zeigen sich Zusammenhänge zwischen den beiden Kompetenzfacetten. Es scheint daher notwendig, weitere Motivationsmaße zu untersuchen, um Erkenntnisse zu Einflussfaktoren auf die Reflexionskompetenz (angehender) Lehrkräfte zu gewinnen.

So liegt ein grundlegendes Problem bei der Erfassung von Reflexion über schriftliche Formate darin, dass Studierende wahrscheinlich nicht alles schreiben, was sie beim Reflektieren auch wirklich denken. Hierzu wird ein kurzer Ausblick auf eine weitere Untersuchung gegeben, im Rahmen derer mittels Lautes-Denken-Interviews die kognitive Validierung der Vignetten im Fokus steht. Da sich das Format von fiktiven Textvignetten in einer Onlinebefragung darüber hinaus evtl. negativ auf die Bearbeitungsmotivation der Studierenden auswirkt – ähnlich einer „didaktisch-induzierten“ Veranlassung von Reflexion in Portfolioleistungen (Häcker, 2019, S. 43) – wurden in weiteren Untersuchungen insbesondere Maße der Testmotivation hinzugezogen, um Testwerte

zur Reflexionskompetenz fundierter interpretieren und damit verbunden die entwickelten Vignetten hinsichtlich ihrer Validität vertieft zu evaluieren.

Qualitätssicherung im bildungswissenschaftlichen Studium – methodologische Überlegungen

Andreas Seifert[1], Christina Watson[2]
[1]Universität Paderborn, [2]SRH Hochschule NRW

Der Test zur Messung Bildungswissenschaftlichen Wissens (BwW; Seifert, Hilligus & Schaper, 2009) wurde mit dem Ziel entwickelt, eine datengestützte Qualitätssicherung und -entwicklung des bildungswissenschaftlichen Studiums in der Lehramtsausbildung zu ermöglichen. Bei der Konstruktion des Kompetenztests stand im Vordergrund, eine handlungsorientierte Testung des bildungswissenschaftlichen Wissens u.a. durch Situational Judgement Tests zu realisieren. Hiermit soll eine fundierte Evaluation der Kompetenzziele der Lehramtsstudiengänge ermöglicht werden. Zudem kann eine differenzierte Analyse nach Inhalts- und Anforderungsdimensionen erfolgen. Die Inhaltsdimensionen differenzieren sich in die Bereiche „Erziehung und Bildung“, „Unterricht und Allgemeine Didaktik“ sowie „Schulentwicklung und Gesellschaft“. Die kognitiven Anforderungen teilen sich in „Wissen reproduzieren, verstehen“, „Reflektieren, Anwenden“ und „Urteilen, Bewerten, Entscheiden“ auf.

In der LEK-Studie (König & Seifert, 2012) wurde eine Kurzversion des Tests an vier bundesdeutschen Universitäten längsschnittlich eingesetzt. Hier konnte eine curriculare Validität des Instruments festgestellt werden, da sich ein deutlicher Leistungszuwachs vom 1. bis zum 4. Semester nachweisen lässt. Auch die Konstruktvalidität des Instruments konnte festgestellt werden, da a) die Binnenstruktur des Instruments mithilfe von strukturanalytischen Methoden bestätigt werden konnte (Seifert & Schaper, 2010) und b) die Zusammenhänge der Bereiche des BwW-Tests zu den unterschiedlichen Facetten des PUW- Tests (König & Blömeke, 2009) sehr plausibel sind (Seifert & König, 2012). Vogelsang und Reinhold setzten dieselbe Kurzversion neben Tests zur Messung fachlicher und fachdidaktischer Kompetenz (Physik) als Prädiktoren für die Güte des Unterrichtshandelns in einer Videographiestudie ein. In einer Stichprobe von 22 angehenden Lehrkräften ergab sich ein signifikanter Zusammenhang des Tests mit fünf von sechs Kriterien guten Unterrichts ($.37 > r > .56$). Somit kann auch von einer kriterialen Validität des Tests ausgegangen werden.

In diesem Beitrag wird die Implementierung einer Online-Version des BwW-Tests im Lehramtsstudium an der Universität Paderborn dargestellt. In der Erhebung ($N = 361$; zw. 2. und 3. Semester) wurde neben der Originalversion mit teilweise offenen Antwortformaten eine Version mit ausschließlich geschlossenen Antwortformaten eingesetzt (die Personen wurden zufällig auf eine der beiden Versionen aufgeteilt). Durch eine IRT-Skalierung ist es gelungen, die unterschiedlichen Versionen des Tests (offene und geschlossene Aufgaben vs. nur geschlossene Aufgaben vs. alle Aufgaben) in drei unterschiedlichen Modellen zu skalieren. Die Version (nur geschlossene Aufgaben) weist eine wesentlich bessere Zuverlässigkeit auf als die Originalversion. Der Zusammenhang zwischen dem Originaltest und dem neuen Test sind ist sehr hoch, $r = .93$, $p < .001$, sodass von einer Äquivalenz ausgegangen werden kann. Auch die Zusammenhänge zwischen den einzelnen Inhaltsfacetten und kognitiven Anforderungsbereichen sind durchweg hoch, $.69 < r < .99$. Ein kritischer Punkt besteht darin, dass die durchschnittliche Schwierigkeit der Items des komplett

geschlossenen Tests signifikant geringer ist als die des Tests mit gemischten Aufgabenformaten, $F(2,283) = 12.817$, $p < .001$. Dies kann zur Folge haben, dass durch Deckeneffekte eventuelle Entwicklungen nicht mehr abgebildet werden können. Neben methodologischen Überlegungen zum Verhältnis von Paper-Pencil- zu Online-Erhebungen soll zusätzlich die inhaltliche Validität eines Kompetenztests nur mit geschlossenen Items thematisiert werden.

Validierung von Testinstrumenten durch Testmotivationsindikatoren

Christina Watson[1], **Andreas Seifert**[2]
[1]SRH Hochschule NRW, [2]Universität Paderborn

Der Validität der Testwertinterpretationen von standardisierten Leistungstests kommt in der Kompetenzforschung eine besondere Bedeutung zu. Insbesondere dann, wenn Testungen ohne direkte Konsequenzen für die Teilnehmenden sind, können Testergebnisse aufgrund mangelnder Motivation verzerrt und somit "a serious threat to the validity of tests" (Baumert & Demmrich, 2001, S. 442) sein. In Anlehnung an Erwartungs-Wert-Modelle kann diese Motivation als domänen- und situationsabhängige Leistungsmotivation verstanden werden (vgl. Eklöf, 2010), für deren Untersuchung es unterschiedliche Möglichkeiten gibt. So können zum einen direkte Testmotivationsmaße erfasst werden wie z.B. die subjektive Einschätzung der Testmotivation (Student Opinion Scale – Thelk et al., 2009). Oder auch indirekte Testmotivationsmaße wie die Erfassung der Testbearbeitungsdauer oder eine "Weiß nicht"-Alternative bei Multiple-Choice Items.

Bei einer Analyse der Testmotivation im Zusammenhang mit der längsschnittlichen Erhebung des BwW-Tests (N = 746) ergaben sich erwartungskonforme Zusammenhänge zwischen Summenscore, "Weiß nicht"-Score und Testbearbeitungsdauer. Weiterhin zeigte sich, dass a) schwierigere Aufgaben besser geeignet sind, Testmotivation zu detektieren und b) Clusteranalysen vier unterschiedliche Typen von Testmotivierten ergeben: 1) stabil motiviert, 2) stabil demotiviert, 3) Anstieg der Motivation und 4) Absinken der Motivation (Watson & Seifert, 2021).

Inwiefern sich ähnliche Typen von Testmotivierten auch für die im Symposium dargestellten Konstrukte und Aufgabentypen identifizieren und daraus verallgemeinerbare Konsequenzen für die Validitätsprüfung ableiten lassen, wird im vierten Beitrag des Symposiums herausgearbeitet.

Hierzu wurden in einer weiteren Studie (N = 349) der Planungstest (Beitrag 1) und der Reflexionstest (Beitrag 2) an zwei Standorten eingesetzt. Der Planungstest umfasst ausschließlich geschlossene Antwortformate (Single Choice), während der Reflexionstest ausschließlich über ein offenes Antwortformat verfügt. Es wurde untersucht, Inwiefern sich das Antwortformat (geschlossen vs. offen) auf die Motivation der Testbearbeitung und somit auf das Testergebnis auswirkt. Die Ergebnisse zeigen, dass das Antwortformat keinerlei Einfluss auf die Motivation hat. Als weiteres direktes Motivationsmaß wurde die Student Opinion Scale (Thelk et al. 2009) eingesetzt. Hier zeigen sich keine signifikanten Zusammenhänge.

Weiterhin wurden indirekte Maße der Testmotivation festgelegt und analysiert. Zum einen die Testbearbeitungsdauer für den geschlossenen Vignettentest. Für den offenen Vignettentest eignete sich die Testbearbeitungsdauer nur bedingt, da diese abhängig von der Tippgeschwindigkeit der Testteilnehmenden ist. Deshalb wurde für den Reflexionstest die Antwortlänge (Zeichen inkl. Leerzeichen) als Motivationsmaß erhoben.

Hierzu wurden Regressionsanalysen durchgeführt. Die Ergebnisse der indirekten Testmotivationsmaße zeigen, dass die Antwortlänge der offenen Vignetten (Reflexionstest) der beste Prädiktor für die Testleistungen in beiden Vignettentests ist. So zeigt sich, dass motivierte Studierende tendenziell in beiden Tests gute Leistungen zeigen.

Die Erfassung von Testmotivationsindikatoren ist für eine valide Testwertinterpretation wichtig, jedoch ist die Auswertung und Interpretation von „motivierten“ und „demotivierten“ Testteilnehmenden schwierig und unterliegt großen Verzerrungstendenzen (rapid guessing vs. solution behavior). Weiterhin sind viele Analysen sehr aufwändig. In diesem Beitrag werden ökonomische individualisierte Maße von Testmotivation vorgestellt und sollen kritisch diskutiert werden.

G-20: Neue Perspektiven auf das Akademische Selbstkonzept

Zeit: Donnerstag, 02.03.2023: 13:15 - 15:00 · *Ort:* R11 T00 D05

Neue Perspektiven auf das Akademische Selbstkonzept

Chair(s): **Christoph Niepel** (Universität Luxemburg, Luxemburg), **Malte Jansen** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin)
*Diskutant*in(nen):* **Ricarda Steinmayr** (TU Dortmund)

Eine Fülle von Forschungsergebnissen hat gezeigt, dass das akademische Selbstkonzept (die Wahrnehmung der eigenen bereichsspezifischen Fähigkeiten von Schüler:innen) einen nachhaltigen Einfluss auf ein breites Spektrum akademischer Ergebnisse hat, wie z. B. schulische Leistungen oder Berufsentscheidungen. Daher verwundert es nicht, dass das akademische Selbstkonzept als eine zentrale Variable schulischer Motivation gilt und Selbstkonzeptforschung sich als eigenes Forschungsfeld etabliert hat. Vor diesem Hintergrund ist Forschung, die zu einem tieferen Verständnis dieses Konstrukts beiträgt, weiterhin von Bedeutung. Das vorliegende Symposium möchte dazu einen Beitrag leisten, indem es vier Beiträge vereint, die neue Perspektiven auf Determinanten und Dynamiken des Selbstkonzepts ermöglichen.

Das akademische Selbstkonzept gilt gemeinhin als ein Konstrukt, welches in der längsschnittlichen Betrachtung überaus stabil ist. Allerdings fehlen Studien, die das akademische Selbstkonzept mittels Experience Sampling im Schulalltag untersuchen und so situative Aspekte des Selbstkonzepts abbilden. Der erste Beitrag geht daher der Frage nach, wie stabil oder variabel Schüler:innen ihre schulischen Fähigkeiten in ihrem Schulalltag tatsächlich wahrnehmen. In einer dreiwöchigen digitalen Tagebuchstudie berichteten Schüler:innen hierzu ihre situativen, fachbezogenen Selbstkonzepte nach jeder Mathematik- und Deutschstunde.

Als zentrale Determinanten des akademischen Selbstkonzepts gelten Effekte sozialer Vergleiche (Vergleiche eigener Leistungen mit den Leistungen der Mitschüler:innen), dimensionale Vergleiche (Vergleiche eigener Leistungen in unterschiedlichen Fächern) und temporale Vergleiche (Vergleiche eigener Leistungen zwischen unterschiedlichen Zeitpunkten).

Im zweiten Beitrag werden soziale Vergleiche in Schulklassen unter Hinzunahme soziometrischer Daten untersucht. Dabei geht er erstens der Frage nach, ob Schüler:innen sich eher mit dem „durchschnittlichen Peer“ oder mit spezifischen Peers (z.B. ihren Freund:innen) vergleichen. Zweitens wird untersucht, ob die Einbindung in soziale Klassennetzwerke (Freundschaft, Unterstützung, Ablehnung) innerhalb einer Klasse beeinflusst, wie stark sich Schüler:innen tatsächlich sozial vergleichen.

Der dritte Beitrag fokussiert auf mögliche negative Effekte dimensionaler Vergleiche auf das akademische Selbstkonzept. Hier wird der Frage nachgegangen, ob die Teilnahme an bilingualen Sachfachunterricht (mit der Unterrichtssprache Englisch) neben positiven Effekten auf das Englisch-Selbstkonzept auch unerwünschte negative Effekte auf das Mathematik-Selbstkonzept mit sich bringt.

Die vierte Arbeit betrachtet soziale, dimensionale und temporale Vergleichseffekte auf das akademische Selbstkonzept integrativ. Alle drei Effekte wurden mit Hilfe eines experimentellen Längsschnittdesigns untersucht, was die Studie klar von bisher vorhandenen Studien abgrenzt. Darüber hinaus wurde geprüft, inwieweit soziale, dimensionale und temporale Vergleiche nicht nur die

Einschätzung der eigenen Fähigkeiten, sondern auch die Einschätzung der Fähigkeiten anderer Personen beeinflussen.

Alle vier Beiträge erweitern somit das Forschungsfeld zur Entstehung und Dynamik des akademischen Selbstkonzepts. Sie greifen dabei innovative, bisher kaum erforschte Forschungsfragen auf und verwenden unterschiedliche, komplementäre Ansätze.

Beiträge des Symposiums

Situative Wahrnehmungen der eigenen Fähigkeiten im Schulalltag: Eine Untersuchung fachbezogener State-Selbstkonzepte mit der Experience Sampling Methode

Christoph Niepel
Universität Luxemburg, Luxemburg

Theoretischer Hintergrund

Zahlreiche Studien haben gezeigt, dass das akademische Selbstkonzept (die mentale Repräsentation der eigenen Fähigkeiten) bemerkenswert stabil über die Zeit ist (siehe z.B. Jansen, 2020; Niepel et al., 2014). Allerdings wurden dabei das akademische Selbstkonzept und seine Stabilität bislang nahezu ausschließlich mittels längsschnittlicher Designs untersucht, die Messintervalle von einigen Monaten bis mehrere Jahre überbrückten (z.B. Möller et al., 2020; Wu et al., 2021). Unser Wissen über kurzfristige Schwankungen und situationsspezifische Ausprägungen des akademischen Selbstkonzepts (State-Selbstkonzept), die Schüler:innen in konkreten Unterrichtssituationen in ihrem Schulalltag tatsächlich erleben, lässt daher viel Raum für Spekulationen. Kürzlich veröffentlichte Studien legen nahe, dass Schüler:innen kurzfristige Schwankungen in ihrem akademischen Selbstkonzept erleben (Hausen et al., 2022; Niepel et al., 2022).

Fragestellung

Das Ziel der vorliegenden Studie war – aufbauend auf diesen ersten Erkenntnissen – die weitere Etablierung des State-Selbstkonzepts im Sinne einer Konstruktvalidierung. Hierfür wurde die Mehrebenenstruktur fachbezogener State-Selbstkonzepte in Mathematik und Deutsch, deren intraindividuelle Variabilität und der Zusammenhang mit theoretisch relevanten Variablen im nomologischen Netz, u.a. dem Trait-Selbstkonzept, untersucht.

Methode

Zu diesem Zweck wurde eine digitale Tagebuchstudie (Experience Sampling Methode via Smartphones) durchgeführt und Daten von insgesamt N = 372 Neunt- und Zehntklässler:innen (18 Klassen) aus 6 verschiedenen Gymnasien in 4 Bundesländern (Baden-Württemberg, Mecklenburg-Vorpommern, Nordrhein-Westfalen und Rheinland-Pfalz) untersucht. Die fachbezogenen State-Selbstkonzepte in Mathematik und Deutsch wurden über einen Zeitraum von 3 Wochen nach jeder Mathematik- bzw. Deutschstunde erhoben (Schüler:innen*Mathematikstunden = 2695; Schüler:innen*Deutschstunden = 2042). Die beiden fachbezogenen State-Selbstkonzepte wurden jeweils mit 3 Items erfragt. Für die Datenanalyse wurden mehrebenenanalytische Strukturgleichungsmodelle in Mplus 8.8 spezifiziert (L1: Situation/Messzeitpunkt, L2: Schüler:in).

Ergebnisse und ihre Bedeutung

Die angenommene Struktur der 3 eingesetzten Items pro State-Konstrukt konnte bestätigt werden. Die beobachteten Reliabilitäten (Omega, auf L1 und L2) der beiden Skalen waren gut. Sowohl das Mathematikselbstkonzept als auch das Deutschselbstkonzept zeigten substanzielle intraindividuelle Variation über die Zeit. Die über die einzelnen Messzeitpunkte aggregierten State-Selbstkonzepte zeigten starke positive Zusammenhänge mit den jeweils korrespondierenden Trait-Selbstkonzepten, die ca. eine Woche vor der dreiwöchigen digitalen Tagebuchstudie im Paper-Pencil-Format erhoben wurden (latente Korrelation von $r = .80$, $p < .01$, für Mathematik und $r = .67$, $p < .01$, für Deutsch).

Weiterhin zeigten sich starke positive Beziehungen zwischen den aggregierten State-Selbstkonzepten und den erhaltenen Noten sowie dem berichteten akademischen Interesse innerhalb desselben Faches, während sich keine statistisch bedeutsamen Zusammenhänge mit den Noten und dem Interesse im jeweils anderen Fach zeigten. Jungen berichteten tendenziell über ein höheres State-Mathematikselbstkonzept als Mädchen, Mädchen tendenziell über ein höheres State-Deutschselbstkonzept als Jungen. Die gefundenen Zusammenhangsmuster der State-Selbstkonzepte für Noten, Gender und Interesse entsprachen damit den für die Trait-Selbstkonzepte gefundenen Zusammenhangsmuster.

Die Befunde dieser Studie implizieren, dass fachbezogene akademische Selbstkonzepte in Mathematik und Deutsch alltäglichen Schwankungen unterliegen. State-Selbstkonzepte können dabei zuverlässig gemessen werden und die gefundenen diskriminanten und konvergenten Zusammenhangsmuster entsprechen denjenigen, die in der vorliegenden und in vorherigen (siehe z.B. Marsh, 2006) Studien für Trait-Selbstkonzepte beobachtet wurden. Die Ergebnisse der vorliegenden Studie können dabei als Ausgangspunkt für weitere Forschungen dienen, um situationsspezifische Schwankungen des akademischen Selbstkonzeptes in konkreten Lernsituationen besser zu verstehen.

Die Rolle sozialer Beziehungen von Schüler:innen für den Big-Fish-Little-Pond-Effekt

Malte Jansen[1], Georg Lorenz[1], Aleksander Kocaj[1], Zsófia Boda[2]
[1]Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin, [2]University of Essex

Theoretischer Hintergrund

Schüler:innen in Klassen mit einem hohen Leistungsniveau zeigen niedrigere Fähigkeitsselbstkonzepte als Schüler:innen mit gleicher individueller Leistung in Klassen mit niedrigerem Leistungsniveau (Big-Fish-Little-Pond-Effekts (BFLPE); Marsh, 1987). Theoretisch wird der BFLPE auf soziale Vergleichsprozesse zurückgeführt– Schüler:innen vergleichen sich mit dem „durchschnittlichen Peer“ (generalized other) um ihre Position in der Leistungsrangreihe der Klasse abzuschätzen (Huguet et al., 2009). Operationalisiert werden Vergleiche mit diesem „durchschnittlichen Peer“ über den Leistungsmittelwert der Klasse. In der Sozialpsychologie wird allerdings davon ausgegangen, dass Schüler:innen sich stärker mit Personen aus ihrer wahrgenommenen In-Group vergleichen („local dominance effect“; Zell et al., 2017). Bisherige Studien zur Rolle von spezifischen Peergruppen innerhalb der Klasse zeigten, dass der Leistungsmittelwert von Freund:innen keinen inkrementellen Effekt auf Fähigkeitsselbstkonzepte über den Klassenmittelwert hinaus hatte („classroom dominance“; Jansen et al., 2022; Koivuhovi et al., 2020; Wouters et al., 2013).

Nur wenige Studien haben allerdings geprüft, ob die Einbindung in soziale Netzwerke in der Klasse beeinflusst, wie stark Schüler:innen sich sozial vergleichen. Schwabe et al. (2019) folgend sollte eine höhere wahrgenommene soziale Einbindung in der Klasse als Puffer gegen negative Einflüsse des BFLPE auf das Fähigkeitsselbstkonzept wirken. Anderseits könnten Schüler:innen, die im Klassennetzwerk eine zentrale Position aufweisen (also mehr soziale Beziehungen haben), sich stärker sozial vergleichen (z.B. weil sie mehr Informationen über die Leistungen ihrer Peers haben; Mussweiler, 2003). Schüler:innen, die von ihren Peers abgelehnt werden, könnten den Wert sozialer Vergleiche niedriger einschätzen (discounting).

Fragestellung

Das erste Ziel der Studie ist eine Replikation erster Studien (Jansen et al., 2022; Koivuhovi et al., 2020; Wouters et al., 2013), die den BFLPE auf Klassenebene Vergleichsprozessen auf der Ebene spezifischer sozialer Beziehungen gegenübergestellt haben. Für das Fach Mathematik wird erwartet, dass auch hier der Leistungsmittelwert der Klasse wichtiger für das Fähigkeitsselbstkonzept ist als der der mittleren Leistungen von Freund:innen (Hypothese 1).

Zweitens sollen mehrere Moderationshypothesen zur Rolle sozialer Beziehungen geprüft werden. Schüler:innen, die sich stärker sozial unterstützt fühlen und besser sozial eingebunden sind, sollten weniger vom BFLPE betroffen sein (Hypothese 2). Schüler:innen, die mehr Informationen über die Leistung ihrer Peers haben (Hypothese 3) oder ein höheres Sozialisationsmotiv haben (Hypothese 4), sollten in stärkerem Maße vom BFLPE betroffen sein. Schließlich sollten Schüler:innen, die von ihren Peers abgelehnt werden (Hypothese 5), sich mit diesen weniger stark vergleichen.

Methode

Wir nutzen Daten der IQB-Bildungstrend 2018 Studie (Stanat et al., 2019, 2022), an der etwa 45.000 Schüler:innen der neunten Jahrgangstufe teilnahmen. Neben einem Kompetenztest im Fach Mathematik sowie Fragen zum Fähigkeitsselbstkonzept und zur wahrgenommenen sozialen Integration wurde auch ein soziometrischer Fragebogen eingesetzt. Dort sollten die Schüler:innen alle Klassenkamerad:innen nominieren, mit denen sie (a) befreundet sind (Freundschaft), (b) die sie bei Problem um Rat fragen (Unterstützung) und (c) neben denen sie nicht sitzen möchten (Ablehnung). Auf dieser Basis wurden soziometrische Indikatoren berechnet, die mit den Moderationshypothesen verknüpft wurden (H2: Outdegree-Zentralität des Unterstützungsnetzwerks und Eingebundenheit im Freundschaftsnetzwerk; H3: Betweenness-Zentralität des Freundschaftsnetzwerks; H4: Outdegree-Zentralität des Freundschaftsnetzwerks; H5: Indegree-Zentralität des Ablehnungsnetzwerks).

Die Auswertung der Daten erfolgte mit Mehrebenen-Strukturgleichungsmodellen unter Verwendung der Software Mplus 7. Zur Schätzung der Moderationseffekte wurden Cross-Level-Interaktionen berechnet.

Erste Ergebnisse & Fazit

Bezüglich des ersten Teilziels konnte die Annahme einer Dominanz des BFLPE auf Klassenebene (H1) bestätigt werden. Die durchschnittliche Mathematikleistung auf Klassenebene zeigte auch nach Kontrolle der individuellen Leistung einen substanziellen negativen Effekt auf das Mathematik-

Selbstkonzept. Die durchschnittliche Leistung der nominierten Freund:innen zeigte keinen inkrementellen Effekt.

Erste Analysen zu den Moderationshypothesen zeigten keine Hinweise darauf, dass die wahrgenommene soziale Eingebundenheit oder die verschiedenen soziometrischen Maße als Moderatoren des BFLPE in Frage kommen. Die Befunde reihen sich somit in eine Reihe von Studien ein, die nahelegen, dass der BFLPE ein sehr robuster Effekt mit nur wenigen Grenzbedingungen zu sein scheint (Seaton et al., 2009).

Effekte bilingualen Sachfachunterrichts (CLIL) auf das akademische Selbstkonzept – Positive Effekte in Englisch aber negative Effekte in Mathematik?

Marlene Wunberg[1], Jürgen Baumert[2], Maja Feddermann[3], Julian F. Lohmann[1], Jens Möller[1]
[1]Christian-Albrechts-Universität zu Kiel, [2]Max-Planck-Institut für Bildungsforschung, Berlin, [3]Institut für Bildungsmonitoring und Qualitätsentwicklung, Hamburg

Theoretischer Hintergrund

Der bilinguale Sachfachunterricht (Content and Language Integrated Learning, CLIL) bezeichnet eine überwiegend in Europa verbreitete Form bilingualen Unterrichts, bei der eine Fremdsprache (L2) als Unterrichtssprache in nicht-sprachlichen Fächern eingesetzt wird (Dalton-Puffer, 2011).

Obwohl die Förderung der Motivation in der L2 ein erklärtes Ziel der CLIL-Implementierung ist (Eurydice, 2006), berücksichtigten Studien zur Effektivität von CLIL bisher selten motivationale Merkmale wie das akademische Selbstkonzept. Angesichts höherer L2-Leistungen von CLIL-Schüler*innen im Vergleich zu nicht-CLIL-Schüler*innen (z.B. Dallinger et al., 2018) ist denkbar, dass das L2-Selbstkonzept von der Teilnahme an CLIL profitiert. Außerdem besitzt CLIL in Deutschland ein hohes Prestige, was sich in Form eines basking-in-reflected glory-Effekts (Marsh, 1984) ebenfalls positiv auf das L2-Selbstkonzept auswirken könnte (Rumlich, 2014). Tatsächlich wiesen CLIL-Schüler*innen in den wenigen vorangegangenen Untersuchungen tendenziell höhere L2-Selbstkonzepte auf als nicht-CLIL-Schüler*innen (z.B. Lo & Lo, 2014; Rumlich, 2017). Die Unterschiede waren jedoch häufig nicht eindeutig auf CLIL, sondern auf Selektionseffekte und verstärkten L2-Unterricht vor Beginn von CLIL (Vorbereitungseffekte) zurückzuführen.

Eine interessante Erweiterung der Untersuchung von CLIL-Effekten auf akademische Selbstkonzepte bietet die Betrachtung des mathematischen Selbstkonzeptes. So verweist die Forschung zur Theorie dimensionaler Vergleiche (Möller & Marsh, 2013) auf negative Auswirkungen von (fremd-)sprachlichen Leistungen auf das mathematische Selbstkonzept (z.B. Möller et al., 2020). Demnach könnten sich die mit CLIL assoziierten höheren L2-Leistungen nachteilig auf das mathematische Selbstkonzept auswirken.

Fragestellung

Wirkt sich die Teilnahme an (englischsprachigem) CLIL positiv auf das Selbstkonzept in Englisch (Forschungsfrage 1), aber negativ auf das Selbstkonzept in Mathematik (Forschungsfrage 2) aus? Wie wirken dabei Selektions- und Vorbereitungseffekte?

Methode

Stichprobe

Datengrundlage war die Hamburger Längsschnittstudie „Kompetenzen und Einstellungen von Schülerinnen und Schülern“ (KESS). Mit den Daten der Klassenstufen 4, 7 und 8 nutzen wir die ersten drei von fünf Messzeitpunkten. Die finale Stichprobe bestand aus 4,841 Schülern*innen aus 257 Grundschulen und 66 Gymnasien (53% weiblich, Alter in Klasse 4: M=10.40), darunter 385 CLIL-Schüler*innen.

Zentrale Variablen

Die Selbstkonzepte in Englisch und Mathematik wurden mit fünf bzw. vier adaptierten Items aus der BIJU-Studie (Baumert et al., 1997) gemessen. Die Teilnahme an CLIL (0 = nein, 1 = ja) wurde anhand eines Items erfasst. Als Maß für Schulleistungen in Englisch und Mathematik wurden Schulnoten und Leistungstests verwendet.

Analysestrategie

Um Selektionseffekte zu kontrollieren, nutzten wir Propensity Score Matching (PSM). CLIL-Effekte auf die Selbstkonzepte schätzten wir mit Strukturgleichungsmodellen. Mittels Difference-in-Differences Analysen untersuchten wir die Entwicklung der Selbstkonzepte auf differenzielle Verläufe zwischen den Gruppen. Fehlende Werte wurden durch multiple Imputation berücksichtigt.

Ergebnisse

Systematische Unterschiede zwischen CLIL- und nicht-CLIL-Schüler*innen in Klasse 4 vor Beginn von CLIL wurden durch PSM erfolgreich ausgeglichen.

Zu Beginn der siebten Klasse wiesen CLIL-Schüler*innen ein signifikant höheres Englisch-Selbstkonzept auf als nicht-CLIL-Schüler*innen ($\beta = 0.11$, $p = .010$). Da CLIL erst in der siebten Klasse begann, ist dieser Effekt mutmaßlich auf Vorbereitungseffekte durch verstärkten Englischunterricht in Klasse 5 und 6 zurückzuführen. Am Ende von Klasse 8 zeigte sich erneut ein signifikanter positiver CLIL-Effekt auf das Englisch-Selbstkonzept ($\beta = 0.13$, $p < .001$), der sich auch unter Berücksichtigung leistungsbezogener Kontrollvariablen als robust erwies ($\beta = 0.06$, $p = .020$). Unter zusätzlicher Kontrolle des Englisch-Selbstkonzeptes aus Klasse 7 war der CLIL-Effekt jedoch nicht länger signifikant. Für das mathematische Selbstkonzept zeigten sich keinerlei signifikante Effekte von CLIL.

Die Difference-in-Differences Analysen ergaben, dass sich die Selbstkonzepte in Englisch und Mathematik zwischen Klasse 7 und 8 bei CLIL und nicht-CLIL-Schüler*innen nicht differentiell entwickelten. CLIL-Schüler*innen konnten ihren signifikanten Vorsprung im Englisch-Selbstkonzept durch CLIL also wahren, aber nicht weiter ausbauen.

Diskussion

Die Teilnahme an CLIL kann dazu beitragen, vorhandene Vorsprünge im Englisch-Selbstkonzept aufrechtzuerhalten und geht gleichzeitig nicht mit negativen Effekten auf das mathematische Selbstkonzept einher. Als ursächlich für die Unterschiede im Englisch-Selbstkonzept zugunsten von CLIL-Schüler*innen sind allerdings Selektions- und Vorbereitungseffekte anzusehen.

Effekte sozialer, dimensionaler und temporaler Vergleiche auf Selbst- und Fremdeinschätzungen bereichsspezifischer Fähigkeiten: Eine experimentelle Längsschnittstudie

Fabian Wolff
Universität Koblenz

Soziale Vergleiche (Vergleiche eigener Leistungen mit den Leistungen anderer), dimensionale Vergleiche (Vergleiche eigener Leistungen in unterschiedlichen Bereichen) und temporale Vergleiche (Vergleiche eigener Leistungen zwischen unterschiedlichen Zeitpunkten) zählen zu zentralen Determinanten bereichsspezifischer Fähigkeitsselbstkonzepte (Trautwein & Möller, 2016). Sie haben gemeinsam, dass sie einen Target (d.h. die Leistung der Person, des Bereichs oder des Zeitpunkts, die/der verglichen wird) mit einem Standard (d.h. die Leistung der Person, des Bereichs oder des Zeitpunkts, mit der/dem verglichen wird) in Beziehung setzen. Außerdem haben sie gemeinsam, dass sie abwärts (wenn der Target besser als der Standard ist), lateral (wenn sich Target und Standard ähneln) oder aufwärts (wenn der Target schlechter als der Standard ist) gerichtet sein können (Wolff et al., 2018).

Bis heute haben zahlreiche Studien gezeigt, dass soziale, dimensionale und temporale Abwärtsvergleiche (Aufwärtsvergleiche) gewöhnlich mit einem erhöhten (verringerten) Selbstkonzept einhergehen (z.B. Wolff & Möller, 2022). Allerdings wurden die meisten dieser Studien im Feld durchgeführt. Bisher gibt es erst drei experimentelle Studien, in denen die Effekte sozialer, dimensionaler und temporaler Vergleiche auf bereichsspezifische Selbstkonzepte simultan untersucht wurden (Wolff et al., 2018, Studie 1; Zell & Strickhouser, 2020, Studie 1 und 2). Interessanterweise zeigte sich nur in einer dieser Studien ein signifikanter temporaler Vergleichseffekt. Dieser Unterschied zwischen den Experimenten und Feldstudien könnte jedoch darauf zurückzuführen sein, dass sich die temporalen Vergleiche in den Experimenten auf Zeiträume von wenigen Minuten bezogen, wohingegen temporale Vergleiche in Feldstudien häufig mehrere Monate oder sogar Jahre umfassten.

Ziel der vorliegenden Studie war es daher, die simultanen Effekte sozialer, dimensionaler und temporaler Vergleiche in einer experimentellen Längsschnittstudie zu untersuchen. Außerdem sollte der Frage nachgegangen werden, inwiefern soziale, dimensionale und temporale Vergleiche die Einschätzungen der Fähigkeiten anderer Personen beeinflussen. Dabei wurde zwischen inferrierten Selbstkonzepten, bei denen Versuchspersonen die Selbstkonzepte anderer Personen erahnen sollten, und Einschätzungen tatsächlicher Fähigkeiten, bei denen Versuchspersonen die Fähigkeiten anderer Personen aus ihrer eigenen Perspektive einschätzen sollten, unterschieden.

Zur Überprüfung der Fragestellungen wurde ein längsschnittliches Experiment mit 411 Studierenden durchgeführt. Diese bearbeiteten an zwei Erhebungszeitpunkten im Abstand von zwei Wochen jeweils einen verbalen und einen figuralen Fähigkeitstest. Am Ende des zweiten Erhebungszeitpunkts erhielten sie manipulierte Rückmeldungen in Bezug auf ihre eigene Leistung und die Leistung eines/einer anderen Studierenden, der/die angeblich ebenfalls am Test teilgenommen hatte. Diese Rückmeldungen (je ein Prozentrang pro Test und Erhebungszeitpunkt) triggerten entweder einen Aufwärtsvergleich oder einen Abwärtsvergleich in Bezug auf eine Vergleichsart, während im Hinblick auf die beiden anderen Vergleichsarten Lateralvergleiche getriggert wurden.

Nach den jeweiligen Rückmeldungen wurden die Versuchspersonen gebeten, ihre Fähigkeiten bzw. die Fähigkeiten des/der anderen Studierenden zur Lösung von Aufgaben des verbalen und figuralen Fähigkeitstests mit Hilfe etablierter Selbstkonzeptskalen einzuschätzen. Die Einschätzungen der Fähigkeiten des/der anderen Studierenden sollten dabei sowohl aus der eigenen Perspektive als auch aus der Perspektive des/der anderen Studierenden vorgenommen werden.

Zur Untersuchung der Vergleichseffekte wurden die mittleren Fähigkeitseinschätzungen zwischen der Aufwärts- und Abwärtsbedingung der jeweiligen Vergleichsart miteinander verglichen. Dabei ergaben die Analysen für alle drei Arten von Fähigkeitseinschätzungen im verbalen und figuralen Bereich signifikante soziale und dimensionale Vergleichseffekte infolge höherer (geringerer) Fähigkeitseinschätzungen nach Abwärtsvergleichen (Aufwärtsvergleichen). Die temporalen Vergleichseffekte fielen jedoch in der Regel nicht statistisch signifikant aus. Die sozialen Vergleichseffekte waren signifikant stärker als die dimensionalen Vergleichseffekte. Ein Vergleich der inferrierten Selbstkonzepte und Einschätzungen der tatsächlichen Fähigkeiten des/der anderen Studierenden ergab insbesondere stärkere soziale Vergleichseffekte auf die Einschätzungen der tatsächlichen Fähigkeiten.

Zusammenfassend stellt die durchgeführte Studie den kausalen Einfluss temporaler Vergleiche bei der Ausbildung von Fähigkeitseinschätzungen infrage. Darüber hinaus zeigt sie, dass sozialen Vergleichen bei der Einschätzung der tatsächlichen Fähigkeiten anderer Personen ein stärkerer Einfluss zukommen dürfte als bei der Einschätzung eigener Fähigkeiten.

Session H

Einzelbeiträge

H-01: Professionswissen von Lehrkräften

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S06 S00 B29
Chair der Sitzung: Caroline Burgwald

Erfassung der professionellen Wahrnehmung zum adaptiven Unterrichten bei Lehramtsstudierenden

Caroline Burgwald, Hardy Ilonca
Goethe Universität in Frankfurt am Main, Deutschland

Als geeignet zum Umgang mit Diversität hat sich der adaptive Unterricht erwiesen (u.a. Hasselhorn & Gold, 2022; Klieme. & Warwas, 2011), der sich durch die Anpassung des Lernangebots an individuelle Lernvoraussetzungen von Lernenden auf der Mikro- und Makroebene des Unterrichts auszeichnet (Bohl et al., 2012; Corno, 2008). Die Anpassungsleistung erfordert von Lehrkräften sowohl Kompetenzen im Bereich der Diagnostik als auch im Bereich der individuellen Förderung, die auch empirisch miteinander zusammenhängen (Beck et al, 2008; Brühwiler, 2017; Franz et al., 2019). Es ist anzunehmen, dass eine Voraussetzung zum adaptiven Unterrichtshandeln auch die Fähigkeit von Lehrpersonen darstellt, im Unterricht selektiv Situationen wahrzunehmen (*noticing*) und diese als lernwirksam bzw. lernhinderlich zu interpretieren (*knowledge-based reasoning*). Diese zwei Prozessebenen gelten als Facetten der professionellen Unterrichtswahrnehmung (Meschede, 2014; Seidel & Stürmer, 2014; Sherin, 2002). Zur Erfassung und Förderung der professionellen Wahrnehmung hat sich die Analyse von videografiertem Unterricht als geeignet erwiesen (Gold et al., 2021).

Obwohl bereits Videovignettentests zu unterschiedlichen Aspekten der Unterrichtsgestaltung vorliegen (z.B. König et al., 2014; Meschede, 2014; Seidel et al., 2010), fehlt bisher ein videobasiertes Instrument zum adaptiven Unterricht. Vor diesem Hintergrund verfolgt die vorliegende Studie zum einen das Ziel, ein Instrument zur Erfassung der professionellen Wahrnehmung von (angehenden) Lehrkräften zu Aspekten adaptiven Unterrichtens zu entwickeln. Zum anderen zielt sie auf die Überprüfung der Validität durch 1. die Untersuchung der Dimensionalität des Instruments sowie 2. die Untersuchung von Zusammenhängen der professionellen Unterrichtswahrnehmung mit Selbstwirksamkeitserwartungen und Einstellungen zu Diversität (Yada et al., 2022).

Testkonstruktion: Das Instrument basiert auf zwei kurzen videografierten Unterrichtssequenzen, die verschiedene Aspekte des adaptiven Unterrichts auf der Mikro- und Makroebene in einem individualisierten Grundschulunterricht (Video 1) und in der Sekundarstufe 1 mit Klassenunterricht und integrierter Gruppenarbeit (Video 2) abbilden. Die zur Beurteilung der Videosequenzen entwickelten vierstufigen Ratingitems beruhen auf der theoretisch wie empirisch nachgewiesenen Konzeptualisierung des Konstrukts des adaptiven Unterrichts. Die insgesamt 32 Items umfassen zum einen Maßnahmen zur lernbegleitenden Diagnostik (z.B. „Die LK stellt Verständnis- und Rückfragen, um den Lernprozess des*der Schüler*in nachzuvollziehen.“) sowie Maßnahmen zur individuellen Förderung (z.B. „Das Feedback der LK bezieht sich auf die fachliche Qualität der von den Schüler*innen gegebenen Antworten."). Die Bearbeitung der Items ist zweischrittig angelegt: Zunächst sollen die Proband*innen feststellen, ob die jeweilige Maßnahme in der gesehenen

Sequenz beobachtbar ist. Sofern dies der Fall ist, sollen sie in einem zweiten Schritt die Passung der Maßnahme in der jeweiligen Unterrichtssequenz bewerten. Durch das zweischrittige Verfahren werden die beiden Prozessebenen der professionellen Wahrnehmung berücksichtigt. Der Prozess der Instrumentenentwicklung umfasste zunächst ein externes Expert*innenrating zur Entwicklung eines Masterratings. Anschließend wurde das Instrument mit N=151 Lehramtsstudierenden unterschiedlicher Fächer und Lehramtsstudiengänge pilotiert und in der Hauptstudie anhand von Strukturgleichungsmodellen überprüft (N=743).

Ergebnisse: Modellvergleiche weisen stichprobenübergreifend auf ein vierdimensionales Konstrukt mit den videospezifischen Faktoren lernbegleitende Diagnostik und individuelle Förderung hin (N=151, CFI =.93; N=743, CFI =.90). Dieser Befund liefert einen ersten Hinweis auf die Validität des Instruments und deckt sich bezüglich eines vorhandenen Videofaktors mit den Ergebnissen der Studie von Meschede (2014). Weiterhin ergeben sich zufriedenstellende bis gute Reliabilitäten, die der Gesamtstichprobe mit N=894 entstammen, für die Skalen "Video 1: lernbegleitende Diagnostik" (8 Items, α=.80), "Video 1: individuelle Förderung" (6 Items, α=.74), "Video 2: lernbegleitende Diagnostik" (10 Items, α=.80) und "Video 2: individuelle Förderung" (8 Items, α=.76). Zusätzliche Untersuchungen zu den Selbstwirksamkeitserwartungen und Einstellungen weisen auf systematische Zusammenhänge mit der professionellen Unterrichtswahrnehmung hin, was weitere wichtige Erkenntnisse zur Konstruktvalidität des Instruments liefert. Das Instrument ermöglicht somit zukünftig die standardisierte Erfassung professioneller Unterrichtswahrnehmung zu Aspekten adaptiven Unterrichtens.

Ein systematisches Review zu TPACK-basierten Professionalisierungsmaßnahmen von (angehenden) Lehrkräften

Armin Fabian, Kenneth Kirchner, Iris Backfisch, Andreas Lachner
Universität Tübingen, Germany

Theoretischer Hintergrund

Um Technologien für einen qualitativ hochwertigen Unterricht einzusetzen, benötigen Lehrkräfte spezifisches technologiebezogenes Fachwissen. Ein bekanntes Modell zur Beschreibung dieses Wissens ist das Modell des technologisch-pädagogischen und inhaltlichen Wissens (TPACK) von Mishra und Koehler (2006). Das TPACK-Modell konzeptualisiert technologiebezogenes professionelles Wissen, d. h. TPACK, als Schnittpunkt von grundlegenden Wissenskomponenten wie Technologischem Wissen (TK), Pädagogischem Wissen (PK) und Inhaltlichem Wissen (CK), sowie weiter als Schnittpunkte von Wissenskomponenten zweiter Ordnung: Technologisch Pädagogisches Wissen (TPK), Technologisch Fachliches Wissen (TCK) und Fachdidaktisches Wissen (PCK). TPACK wird als zentrale Voraussetzung für die erfolgreiche Integration von Technologien in den Fachunterricht angesehen.

Dank seiner Eingänglichkeit wurde das TPACK-Modell häufig als Grundlage für die Gestaltung von adäquaten Lehreraus- und fortbildungen verwendet (z. B. Voogt et al., 2013). Jüngste Untersuchungen Forschung hatben jedoch eine Vielzahl unterschiedlicher Ansätze zur Förderung des Professionswissens gefundenzum Vorschein gebracht, die im TPACK-Modell verankert sind (Wang et

al., 2018, Willermark, 2018). Die Verschiedenheit dieser Ansätze macht es allerdings schwierig, entsprechende Forschungsergebnisse und Fortbildungskonzepte miteinander zu vergleichen.

Vor diesem Hintergrund haben wir ein systematisches Review durchgeführt, in der wir TPACK-basierte Aus- und Fortbildungen (veröffentlicht zwischen 2006 und 2020) systematisch analysierten. Um dies zu tun, haben wir mittels qualitativer Inhaltsanalyse (Mayring, 2014) untersucht, (F1) welche TPACK-Komponenten in den jeweiligen Fortbildungen primär fokussiert wurden. Zusätzlich haben wir (F2) systematisch untersucht, welche Kombinationen von TPACK-Komponenten häufig zusammen adressiert wurden, um verschiedene Fortbildungskonzeptionen zu systematisieren und so Ergebnisse für Forschung und Praxis vergleichbarer zu machen.

Methode

Angelehnt an das PRISMA-Framework (Moher et al., 2019) haben wir einen umfassenden Suchterm verwendet, um aus sechs Datenbanken Literatur zu extrahieren (N = 2907). Die extrahierten Beiträge wurden von zwei Ratern unabhängig nach vorab festgelegten Einschlusskriterien gescreent, was in einem finalen Datensatz von N = 140 Beiträgen resultierte. Anhand eines Kodierschemas wurde für jede Fortbildung entschieden, welche TPACK-Komponente explizit adressiert wurde, und welche nicht (multiples Kodieren). Um die Qualität des Kodierungsschemas sicherzustellen, wurden 20 % der Beiträge blind kodiert (κ = .75). Für die nachfolgende Analyse wurden φ-Korrelationen berechnet und χ2-Tests durchgeführt, um den Zusammenhang des Auftretens zweier TPACK-Komponenten zu untersuchen.

Egebnisse

TK war die am häufigsten fokussierte TPACK-Komponente (80,7 %) und CK die am seltensten adressierte TPACK-Komponente (35,0 %) in den inkludierten Fortbildungen. PK wurde in 39.2 % der Fälle adressiert. Die Wissenskomponenten zweiter Ordnung wurden jeweils in rund der Hälfte der Fortbildungen adressiert (TPK: 50,7 %; TCK: 43,4 % und PCK: 50,0 %). TPACK als zentrales Konstrukt wurde in 76,4 % der Fälle explizit gefördert. Hinsichtlich des gemeinsamen Auftretens zweier TPACK-Komponenten zeigen die Ergebnisse einen hohen positiven signifikanten Zusammenhang zwischen PCK und TPACK ($\varphi_{PCK,TPACK}$ = .19*) und zwischen TCK und TPACK ($\varphi_{TCK,TPACK}$ = .32**). Für TK lassen sich keine signifikanten Zusammenhänge mit anderen TPACK-Komponenten feststellen (-.15 < φ < .14). CK hingegen steht in positivem Zusammenhang mit den anderen C-Komponenten, wie PCK ($\varphi_{CK,PCK}$ = .23*), TCK ($\varphi_{CK,TCK}$ = .34**) und TPACK ($\varphi_{CK,TPACK}$ = .19**).

Diskussion

Der Fokus auf TK-bezogene Instruktionen in Fortdbildungen deutet daraufhin, dass Forschende TK als wesentliche Voraussetzung von technologiebezogenem Professionswissen verstehen. Gleichzeitig zeigen die nicht signifikanten Zusammenhänge von TK mit den anderen TPACK-Komponenten, dass insbesondere digitalisierungsbezogene Bedienfertigkeiten über die Fortbildungen hinweg unsystematisch adressiert wurden. Die positiven Zusammenhänge zwischen CK, PCK, TCK und TPACK weisen daraufhin, dass sich TPACK häufig fachspezifisch angenähert wird, was im Einklang mit Forderungen aus vergangenen Forschungsprojekten ist (Voogt et al.,

2013). Insgesamt bestätigen die empirischen Ergebnisse des Reviews das Gefühl der Diversität von TPACK-basierten Fortbildungskonzepten. Weitere Forschung ist allerdings erforderlich, um jene Fortbildungen umfänglicher zu systematisieren. Hier wäre ein vielversprechender Ausgangspunkt, die Fortbildungen genauer daraufhin zu untersuchen, wie die verschiedenen TPACK-Komponenten miteinander verwoben und kombiniert wurden, um Professionswissen zum Einsatz digitaler Medien zu fördern.

Das Professionswissen von (angehenden) Lehrkräften im Fach Sozialwissenschaften: Entwicklung und Validierung eines Testinstruments

Dorothee Gronostay[1], **Sabine Manzel**[2]
[1]TU Dortmund, Deutschland; [2]Uni Duisburg-Essen, Deutschland

Als theoretische Grundlage wird auf das generische Modell professioneller Handlungskompetenz von Lehrpersonen nach Baumert und Kunter (2006) rekurriert. Die angenommene Dreiteilung des Professionswissens in die Bereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen geht auf Arbeiten von Shulman (1986, 1987) zurück und hat sich in der Lehrerprofessionsforschung etabliert. Eine fachspezifische Modellierung entlang des Modells wurde erstmals durch Weschenfelder (2014) im Forschungsprojekt „Professionelle Kompetenz von Politiklehrkräften“ (Weißeno, Weschenfelder & Oberle, 2013) für die Domäne der Politik vorgelegt. Ein Testinstrument, welches alle drei Fachwissensbereiche des Lehramtsstudienfachs SoWi und die Fachdidaktik gleichermaßen berücksichtigt, fehlt im Kontext der Professionsforschung im Fach jedoch.

Das Ziel dieser Studie besteht daher in der Entwicklung und Validierung eines standardisierten und ökonomisch einsetzbaren Testinstruments (Bearbeitungszeit: max. 60 Min.) zur Erhebung des Professionswissen von (angehenden) Lehrkräften im Lehramtsstudienfach SoWi. Ausgangspunkt sind Vorarbeiten im Kontext des Teilprojekts Sozialwissenschaften im BMBF-geförderten Projekt „ProViel“ der Univ. Duisburg-Essen. Der SoWis-L soll deklaratives Professionswissens im Fach SoWi in den fachdidaktischen und fachwissenschaftlichen Inhaltsbereichen auf Reproduktions- und Anwendungsniveau erfassen. Die Forschungsfragen beziehen sich auf strukturelle (Reliabilität und Dimensionalitätsprüfung) und auf externale Validitätsaspekte (Methode der bekannten Gruppen, Korrelation mit Außenkriterien).

Die Entwicklung der 134 Testaufgaben erfolgte systematisch entlang der Dimension „Wissensbereich“ (mit den Ausprägungen Politik, Wirtschaft, Soziologie und Fachdidaktik), „Facette“ (drei bis vier Ausprägungen je Wissensbereich) und „kognitive Prozesse“ (mit den Ausprägungen „erinnern und abrufen“ und „verstehen und anwenden“ (Anderson et al., 2001, Großschedl et al., 2015; Kauertz et al., 2010). Auf Basis einer Pilotierungsstudie an N = 229 Lehramtsstudierenden wurden 63 Items für die Hauptstudie ausgewählt.

Als Datengrundlage der Hauptstudie diente eine Stichprobe von *N* = 374 Lehramtsstudierenden des Fachs SoWi der Universität Duisburg-Essen. Die Papier-Bleistift-Erhebung erfolgte über vier Semester (SoSe 18 bis WiSe 19/20) in Pflichtveranstaltungen des Lehramtsstudiums SoWi.

Die Reliabilitäten der Politik-, Wirtschafts- und Fachdidaktikskala lagen in einem zufriedenstellenden Bereich, für die entwickelte Soziologie-Skala galt dies jedoch nicht (so dass diese von den

weitergehenden Analysen ausgeschlossen wurde). Bei der Prüfung der Struktur des 46-Items-Instruments führte das dreifaktorielle Modell sowohl gegenüber dem Generalfaktormodell ($\Delta \chi^2$ = 93,38, df = 5, p < .001) als auch gegenüber dem zweifaktoriellen Modell ($\Delta \chi^2$ = 46,95, df = 3, p < .001) zu einer signifikanten Modellverbesserung.

Die befragten GyGe-Studierenden erreichten in allen Wissensbereichen durchschnittlich bessere Testleistungen als die HRG-Studierenden. Dies gilt für die Bereiche Politik (t_{347} = 4,42, p < .001, d = ,57), Wirtschaft (t_{347} = 3,68, p < .001, d = ,48) und Fachdidaktik (t_{347} = 1,69, p = .046, d = ,22). Diese Ergebnisse sind in hohem Maße erwartungskonform. Ebenso steigt die Testleistung Politik und Fachdidaktik zwischen BA- und MA-Studium an. Eine Kontraststichprobe von zwölf befragten Lehrkräften im Schuldienst erreichte darüber hinaus eine signifikant höhere Testleistung in allen Wissensbereichen, was als Hinweis auf eine Sensitivität des Testinstruments für Expertiseunterschiede verstanden werden kann. Weiter sind die Testleistungen im SoWis-L in allen drei Wissensbereichen signifikant negativ mit der Abiturnote verknüpft, am stärksten im Bereich des wirtschaftlichen Wissens (r = -,27, p < .001), am geringsten im Bereich des fachdidaktischen Wissens (r = -,16, p = .004). Dieses Ergebnis ist hypothesenkonform.

Der SoWis-L ist das bislang einzig standardisierte Testinstrument für den Bereich des Professionswissens im Lehramt SoWi, das den Integrationscharakter des Fachs abbildet. Als innovativ kann gelten, dass der politische und wirtschaftliche Wissensbereich sowie die Fachdidaktik getrennt modellierbar sind und so Forschungsfragen zum Professionswissen von Lehrkräften mit höherem Differenzierungsgrad untersuchbar sind als bislang. Ein Desiderat besteht weiterhin in der Entwicklung einer reliablen Soziologieskala. Insgesamt sind weitere Analysen an größeren Stichproben erforderlich, um die Validität des Instruments bezüglich der intendierten Interpretation und Nutzung umfänglich abschätzen zu können (vgl. AERA et al., 2014).

Wie beurteilen Fach- und Schulleitungen angehende Lehrer:innen im Referendariat? Erkenntnisse aus einer qualitativ-rekonstruktiven Dokumentenanalyse von schriftlichen Gutachten.

Christoph Kruse
Westfälische Wilhelms-Universität Münster, Deutschland

Theoretischer Hintergrund und Fragestellungen

Vor dem Hintergrund der Debatte um die Professionalität von Lehrer:innen und aufgrund der immanenten Selektionsthematik erscheinen Beurteilungen in der Lehrer:innenbildung als gleichsam kaum erforschter und dennoch hoch relevanter Forschungsgegenstand (vgl. z.B. König, 2020; Rauin, 2014). Dies ist besonders im Rahmen des Vorbereitungsdienstes (VD) interessant, da hier ein breites Spektrum des Lehrer:innenhandelns zur Beurteilung steht. Entsprechend formulieren Reintjes, Bellenberg und Labott (2021) ein klares Desiderat bezüglich „fachlicher Erwartungshaltungen“ (ebd., S. 128) sowie zu „Erwartungshaltungen von [...] beurteilenden Schulleitungen, die als wichtige Akteure im VD fungieren“ (ebd.).

Im Beitrag wird ein Forschungsprojekt vorgestellt, in dem als Zugang zu diesen Erwartungen zentrale Beurteilungsdokumente aus Nordrhein-Westfalen (NRW) analysiert werden: Am Ende des VD verfassen Fach- und Schulleitungen Gutachten, in denen Fähigkeiten und vollzogene Entwicklungen

von Lehrer:innen im Vorbereitungsdienst (LiV) dargelegt, begründet und letztendlich benotet werden. Konzepte zur Professionalität und Professionalisierung von Lehrer:innen stellen den Bezugspunkt des Projekts dar. Schließlich wird im Professionalitätsdiskurs verhandelt, was von Lehrpersonen als professionelle Praxis erwartet werden kann. Statt sich einem spezifischen, erkenntnisleitenden Professionalitätsansatz anzuschließen (vgl. Kalthoff, 2019), werden in diesem Projekt zentrale Konzepte der Professionalität jeweils hinsichtlich ihrer normativen Dimension expliziert und zur theoriebasierten Diskussion der Erkenntnisse genutzt (vgl. Fragestellung 3.). Insofern wird empirisch anhand der Gutachten der Frage nachgegangen, „wie Professionalität dargestellt, thematisiert oder verhandelt wird“ (Bennewitz, 2020, S. 192).

Unter dem übergeordneten Forschungsinteresse, wie Fach- und Schulleitungen angehende Lehrer:innen im Referendariat in NRW beurteilen, wurden daraus die folgenden *Fragestellungen* materialbasiert entwickelt:

1. Welche Erwartungen an eine professionelle Lehrperson am Ende des Referendariats werden in den Gutachten deutlich?
2. Wie werden vorhandene und fehlende Fähigkeiten in der Beurteilung erklärt und die Beurteilung begründet?
3. In welchem Verhältnis steht das unter I. und II. rekonstruierte Gesamtbild einer professionellen Lehrperson zu den im wissenschaftlichen Diskurs verhandelten Professionalitätsansätzen?

Methodisches Vorgehen

Mit Unterstützung der zuständigen Landesbehörde konnte kriteriengeleitet ein anonymisierter *Datensatz* von 60 Gutachten zusammengestellt werden, anhand dessen eine möglichst große Varianz der schriftlichen Beurteilungspraxis in den Blick genommen werden kann. Zur Auswertung wird sich im Rahmen einer explorativen Dokumentenanalyse (Hoffmann, 2018) auf die *Grounded Theory Methodologie* (GTM, Strauss, 1998) bezogen. Notwendig für ein konstant vergleichendes Vorgehen gemäß des theoretical samplings der GTM ist folglich eine zweite Stufe der Fallauswahl *innerhalb* des o.g. Gesamtkorpus. Die GTM mit dem inhärent induktiven und stark gegenstandsangemessenen Vorgehen scheint insbesondere für die Erforschung des unerforschten und spezifischen Gegenstands geeignet. Ferner ermöglicht diese Methodenwahl die umfassende rekonstruktive Analyse zahlreicher Gutachten als auch die systematische Berücksichtigung des Entstehungskontexts der Gutachten.

Ergebnisse

Zunächst ist bedeutsam, dass in der formalen Gutachtenvorlage als Beurteilungskriterien die (bildungswissenschaftlichen) Handlungsfelder und Kompetenzen gemäß der „Standards für die Lehrerbildung“ (KMK, 2019) angeführt werden. Rekonstruktionen der Gutachtentexte deuten darauf hin, dass die Anbindung an diese geforderten Beurteilungskriterien in unterschiedlicher Weise geschieht: Einerseits scheint eine Perspektive auf, die den Lehrer*innenberuf als *erlernbar* im Sinne der Kompetenzformulierungen beschreibt. Andererseits wird auf Aspekte einer „Lehrerpersönlichkeit“ rekurriert, was eine Erlernbarkeit des Berufs eher infrage stellt. Dass in den Gutachten beschriebene Entwicklungen der LiV vornehmlich als Defizitkonstruktion angeführt werden, stützt die Vermutung ein Vollbild an Kompetenzen stünde als grundlegende Erwartung im Hintergrund der Beurteilung. Das lässt sich im Widerspruch zur Annahme von „Professionalität [...] als berufsbiografische[m] Entwicklungsproblem“ (Terhart, 2011, S. 208) diskutieren. Zudem deutet sich eine Beurteilungspraxis

an, die die vorzeitige Schließung etwaiger (pädagogischer) Ungewissheiten begünstigt. Die Erkenntnisse können also sowohl aus einer kompetenztheoretischen als auch berufsbiografischen und strukturtheoretischen Perspektive auf Professionalität (vgl. Fragestellung 3.) problematisiert werden.

Insgesamt lassen sich somit im Beitrag empirische Erkenntnisse zur Diskussion um die Problematik des „unzureichenden Konsens über die Erfolgskriterien einer gelungenen Berufseinführungsphase im wissenschaftlichen und politischen Diskurs“ (Rauin, 2014, S. 575) erwarten.

H-02: Motivation, Selbstkonzept, Selbstwirklichkeit, leistungszielorientiert und als Vergleich

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S06 S00 B32
Chair der Sitzung: Rebecca Schneider

Selbstkonzepte mediieren die Beziehung zwischen fachspezifischen Leistungen und intrinsischen Werten

Rebecca Schneider[1], **Fabian Wolff**[2]
[1]Institut zur Qualitätsentwicklung im Bildungswesen e. V. (IQB), Germany; [2]Universität Koblenz-Landau

Einflüsse sozialer und dimensionaler Vergleiche von schulischen Leistungen auf die Ausbildung fachspezifischer intrinsischer Werte (auch Interessen) sind für Sekundarschüler*innen empirisch gut belegt (z.B. Gaspard et al., 2018). Verschiedenste Theorien, u.a. die Erwartungs-Wert-Theorie (Wigfield & Eccles, 2002), die Self-Determination Theory (Deci & Ryan, 2000) sowie die Person-Gegenstands-Theorie des Interesses (Krapp, 2005), nehmen zudem eine Wirkung fachspezifischer Selbstkonzepte auf intrinsische Werte/Interessen an. Entsprechend deuten erste Studien auf eine mediierende Wirkung von Selbstkonzepten auf den Leistungs-Wert-Zusammenhang hin. Diese Studien sind bisher jedoch auf zwei Schulfächer (z. B. Götz et al., 2008; Nagy et al., 2008) oder ausschließlich mathematisch-naturwissenschaftliche Fächer (Guo et al., 2017) begrenzt. Dies könnte jedoch insofern problematisch sein, als das Leistungen in verschiedenen Fächern an der Bildung eines fachbezogenen akademischen Selbstkonzepts beteiligt sind und insbesondere die Beschränkung eines Modells auf nur zwei Fächer oder auf mathematisch-naturwissenschaftliche oder sprachliche Fächer zu verzerrten Ergebnissen führen kann (siehe Möller et al., 2020, für eine Diskussion). Ziel des Beitrags ist daher eine Erweiterung bisheriger Forschung zur Mediationshypothese, für welche wir die Beziehungen zwischen fachspezifischen schulischen Leistungen, Selbstkonzepten und intrinsischen Werten in sechs Schulfächern untersucht haben.

Die Analysen basieren auf Daten von N = 1762 Schüler*innen aus 111 Klassen der neunten bis zwölften Klassenstufe. Erfasst wurden die Noten der Schüler*innen in Mathematik, Physik, Wirtschaft/Politik, Deutsch, Englisch und der zweiten Fremdsprache als Leistungsindikatoren. Diese wurden umkodiert, sodass höhere Werte besseren Leistungen entsprechen. Die fachspezifischen Selbstkonzepte und intrinsischen Werte wurden mit je 3, über die Fächer hinweg parallel formulierten Items erfasst (z. B. „Ich bin gut in [Fach]“, „Ich mag [Fach]“; 6-stufige Antwortskala von „trifft überhaupt nicht zu“ bis „trifft völlig zu“). Mithilfe von Strukturgleichungsmodellen in Mplus wurden die Noten der Schüler*innen mit den intrinsischen Werten in diesen Fächern in unterschiedlich spezifizierten Modellen ([a] MLR-Schätzer und Berücksichtigung der hierarchischen Datenstruktur, [b] ML-Schätzer und bootstrapping) in Beziehung gesetzt, die fachspezifischen Selbstkonzepte wurden dabei jeweils als Mediatoren aufgenommen.

Die Ergebnisse der Strukturgleichungsmodelle zeigen konsistent eine vollständige Mediation der Beziehung zwischen Noten und intrinsischen Werten durch die Selbstkonzepte. Die direkten Pfade von den Noten zu den entsprechenden intrinsischen Werten waren i. d. R. nicht signifikant. Die wenigen signifikanten Koeffizienten der direkten Pfade waren zumeist unterhalb der Grenze kleiner Effekte ($-.13 \leq \beta \leq -.07$).

Bezogen auf die indirekten Pfade von den Noten zu den entsprechenden intrinsischen Werten vermittelt durch Selbstkonzepte, waren die Koeffizienten für Pfade mit korrespondierender Note,

Selbstkonzept und intrinsischem Wert am stärksten ausgeprägt (.54 ≤ β ≤ .77), was die Bedeutung sozialer Vergleiche für die Bildung von Selbstkonzepten bzw. intrinsischen Werten unterstützt. Darüber hinaus wurden z. T. substantielle indirekte Pfadkoeffizienten von der Leistung auf die intrinsischen Werte gefunden, wenn sich das Selbstkonzept und der intrinsische Wert auf das gleiche Schulfach bezogen (z. B. von der Deutschnote auf den intrinsischen Wert in Physik vermittelt durch das Physikselbstkonzept; -.19 ≤ β ≤ .11).

Zusätzliche Analysen zeigen zudem eine vergleichbare Passung zwischen dem Mediationsmodell mit direkten Pfaden von den Noten auf die intrinsischen Werte und einem Mediationsmodell ohne diese direkten Pfade (ΔCFI < .01, ΔRMSEA < .015).

Insgesamt deuten die Ergebnisse damit auf einen zweistufigen Prozess hin: Während fachspezifische Selbstkonzepte durch Vergleichsprozesse fachspezifischer Leistungen gebildet werden (vgl. auch Annahmen zum Internal/external-frame-of-reference-Modell; Marsh, 1986), sind diese fachspezifischen Selbstkonzepte die primäre Quelle für die Bildung fachspezifischer intrinsischer Werte. Unsere Befunde sollen in die bisherige Literatur zur Genese fachspezifischer Selbstkonzepte und Werte eingeordnet werden. Zudem soll diskutiert werden, welche Bedeutung Selbstkonzepte für Beziehungen zwischen Leistung und weiteren Wertfacetten der Erwartungs-Wert-Theorie (z.B. Wichtigkeit) haben könnten.

Häufigkeiten und Auswirkungen akademischer Vergleiche im Alltag

Alexandra Petrak[1,2], Jens Möller[1], Friederike Helm[1], Fabian Wolff[2]
[1]Institut für Pädagogisch-Psychologische Lehr- und Lernforschung, Christian-Albrechts-Universität zu Kiel; [2]Institut für Psychologie, Universität Koblenz-Landau, Campus Koblenz

Theoretischer Hintergrund

Studierende können ihre Leistungen in einem Bereich mit denen ihrer Kommiliton*innen (sozialer Vergleich), im Zeitverlauf (temporaler Vergleich) oder ihren Leistungen in einem anderen Bereich (dimensionaler Vergleich) vergleichen (Trautwein & Möller, 2016). Solche Vergleichsprozesse wirken sich auf akademischen Selbstüberzeugungen und Leistungen aus. Bisher ist jedoch wenig dazu bekannt, mit welcher Häufigkeit die einzelnen Vergleiche im Alltag auftreten und wie sie sich auf das Selbstkonzept auswirken.

Tagebuchstudien konnten bereits zeigen, dass soziale und dimensionale Vergleiche im Alltag mit Änderungen von Stimmungszuständen und motivationalen Variablen einhergehen (Baumann, 2020; Möller & Husemann, 2006; Olson & Evans, 1999; Wheeler & Miyake, 1992). Allerdings gibt es bisher weder Tagebuchstudien über temporale Vergleiche, noch zu den Effekten der Vergleiche auf akademische Selbstkonzepte.

Die vorliegende Studie untersucht deshalb erstmals gleichzeitig mehrere Vergleichsarten und ihre Folgen. Konkret werden die Häufigkeit des Auftretens sozialer, temporaler und dimensionaler Vergleiche im akademischen Kontext sowie deren Auswirkungen auf das Selbstkonzept untersucht.

Zur Auftretenshäufigkeit gibt es Hinweise, dass Studierende mehr soziale (Olsen & Evans, 1999; Wheeler & Miyake, 1992) als dimensionale (Baumann, 2020; Möller & Husemann, 2006) und mehr soziale als temporale Vergleiche durchführen (Wilson & Ross, 2000).

Zur Vergleichsrichtung wurde eine allgemeine Präferenz für Aufwärtsvergleiche (d.h. mit besseren Leistungen) für soziale (Wilson & Ross, 2000) und dimensionale Vergleiche (Baumann, 2020; Möller & Husemann, 2006) gefunden, für temporale Vergleiche eine Präferenz für Abwärtsvergleiche (d.h. mit einem niedrigeren Vergleichsstandard) mit der Vergangenheit (Wilson & Ross, 2000).

Vergleichsprozesse beeinflussen Selbstkonzepte in verschiedenen Bereichen (Möller et al., 2011, 2020). Experimentelle Studien fanden, dass soziale, temporale und dimensionale Abwärtsvergleiche zu einer Verbesserung, Aufwärtsvergleiche zu einer Verschlechterung des Selbstkonzepts führen (Möller & Köller, 2001; Wolff et al., 2018).

Hypothesen

Wir nahmen an,

(H1) dass Studierende mehr soziale und temporale als dimensionale Vergleiche durchführen,

(H2) soziale und dimensionale Vergleiche überwiegend aufwärts-, und temporale Vergleiche abwärtsgerichtet sind,

(H3) das bereichsspezifische Selbstkonzept nach Abwärtsvergleichen (Aufwärtsvergleichen) höher (niedriger) als vor dem Vergleich ist.

Methode

Für die vorliegende Studie wurden N = 131 Studierende gebeten, ihre alltäglichen akademischen Vergleiche innerhalb einer Woche in einem Online-Tagebuch festzuhalten und dabei Inhalt, Richtung und Auswirkung eines jeden Vergleichs auf das Selbstkonzept zu beschreiben.

Zur Messung der Veränderung des Selbstkonzepts wurden die Teilnehmenden gebeten, ihre Fähigkeit, ihr Talent und ihre Leistung im Ausgangsbereich (was verglichen wurde) nach dem Vergleich im Vergleich zu vor dem Vergleich einzuschätzen (7-Punkte-Likert-Skala von 1= *viel schlechter* bis 7= *viel besser*, α= .86).

Mit χ^2-Ähnlichkeitstests wurde untersucht, ob sich das Auftreten der Vergleiche signifikant voneinander unterschied. Mithilfe eines Regressionsmodells wurde der Vergleichseffekt auf das Selbstkonzept untersucht. Hierbei wurden die Vergleiche (z.B. temporaler Aufwärtsvergleich) als Dummyvariablen kodiert und die genestete Struktur von Mehrfachmessungen pro Person beachtet. Vergleichseffekte wurden mittels Differenzbildung mit dem jeweiligen Lateraleffekt berechnet (*Mplus*-Befehl: MODEL CONSTRAINT).

Ergebnisse

(H1) Wie erwartet, traten soziale und temporale Vergleiche häufiger auf als dimensionale Vergleiche, $\chi^2(2)= 297.47$, $p < .001$.

(H2) Soziale Vergleiche waren überwiegend aufwärtsgerichtet, $\chi^2(2)= 15.02$, $p = .001$, temporale Vergleiche abwärtsgerichtet, $\chi^2(2)= 22.71$, $p < .001$. Für dimensionale Vergleiche zeigte sich kein Unterschied in der Häufigkeit der Vergleichsrichtung.

(H3) Fast alle Vergleiche zeigten einen Effekt auf das Selbstkonzept in der Ausgangsdomäne: Nach sozialen ($\Delta\beta$= .33, p < .001) und temporalen ($\Delta\beta$= .36, p < .001) Abwärtsvergleichen nahm es zu, nach sozialen ($\Delta\beta$= –.20, p < .001) und dimensionalen ($\Delta\beta$= –.12, p < .001) Aufwärtsvergleichen ab.

Diskussion

Unsere Studie ist die erste, die mehrere Vergleichsprozesse im Alltag innerhalb einer einzigen Tagebuchstudie untersucht. Unsere Ergebnisse unterstreichen, dass soziale, temporale und dimensionale Vergleiche im Alltag von Studierenden vorkommen und Auswirkungen auf das Selbstkonzept haben. Ihre weitere Erforschung erscheint vielversprechend für das Verständnis der Selbstkonzeptentwicklung und als Grundlage für die Entwicklung einer allgemeinen Theorie akademischer Vergleichsprozesse.

Kann die schulbezogene Selbstwirksamkeitserwartung von Schüler*innen mittels weiser Interventionen gestärkt werden? Ergebnisse dreier explorativer, quasi-experimenteller, kontextsensibler Interventionsstudien.

Zuzanna M. Preusche, Kerstin Göbel
Universität Duisburg-Essen, Germany

Im Zuge der Adoleszenz lässt sich bei einem Großteil der Schüler*innen ein systematisches schulisches Disengagement beobachten, dessen Folgen sich in Schulleistungen bzw. Qualifikationen widerspiegeln (Wang et al., 2019). Die Auseinandersetzung mit Gründen für dieses Phänomen und der daraus resultierende Fokus auf die dominierenden Disparitäten zwischen bestimmten Schüler*innengruppen offenbaren die (interdisziplinäre) Komplexität der Erklärungen (Radtke & Gomolla, 2009; Göbel & Buchwald, 2017). Basierend auf unterschiedlichsten Faktoren, Dispositionen oder Erfahrungen, können Schüler*innen z. B. ein geschwächtes schulisches Wohlbefinden, mangelndes Zugehörigkeitsgefühl oder niedriges akademisches Selbstkonzept entwickeln (z. B. Wang et al., 2019; Citivillo et al., 2021; Hascher, 2004; Göbel & Preusche, 2022). Eine Aufgabe für die Unterrichtsentwicklung kann daher in der Stärkung der Schüler*innen sowie in dem Ausgleich möglicher Disparitäten gesehen werden. Dieses Ziel verfolgen neben anderen Maßnahmen auch sog. wise Interventions, welche bisher vor allem im US-amerikanischen Raum genutzt werden, um Schüler*innen zu stärken. Trotz der mittlerweile deutlich gewachsenen empirischen Befundlage (z. B. Wu et al., 2021) und der ersten Implementationsvorschläge auch im deutschsprachigen Raum (SVR, 2020), können die Effekte der ersten Studien teilweise nicht repliziert werden, weshalb viele Autor*innen für die systematische Auseinandersetzung mit den Wirkungsmechanismen sowie den Implementationsmöglichkeiten plädieren (u. a. Bryan et al., 2021). Zu diesem Forschungsdesiderat möchte die vorliegende Studie einen Beitrag leisten, um die Implementationsmöglichkeiten dieser kurzen, jedoch vielversprechenden Interventionen für den deutschsprachigen Raum weiter zu eruieren. Die vorliegende quasi-experimentelle Studie hatte zum Ziel, die drei Interventionsdesigns Value Affirmation (n = 103), Social Belonging (n = 87) und Growth Mind Set (n = 78; sowie einer delayed treatment Kontrollgruppe, n = 79), explorativ und kontextsensibel, geleitetet durch die jeweiligen Klassenlehrpersonen mit Schüler*innen der 8. und 9. Klassen in Gesamtschulen in NRW durchzuführen und zu prüfen, ob die Designs einen Effekt auf die schulische Selbstwirksamkeitserwartung (α = .88; 4 Items; adapt. Dickhäuser et al., 2002), welche als Proxy für

das Selbstkonzept und Schulleistungen angesehen wird, haben können. In der Value Affirmation-Kondition reflektieren die Schüler*innen über für sie wichtige Werte um ein stärkeres, generelles Selbstkonzept zu entwickeln; die Social Belonging-Kondition vermittelt den Schüler*innen soziale (schulische) Zugehörigkeit als einen universalen Prozess; die Growth Mind Set-Kondition lehrt die Schüler*innen eine wachstumsorientierte Grundhaltung gegenüber den eigenen schulischen Fähigkeiten. Eine Mixed ANOVA wurde durchgeführt, um zu prüfen, inwiefern sich die drei Experimentalgruppen und die Kontrollgruppe zwischen dem Pre- und Posttest unterscheiden. Nach Prüfung der Voraussetzungen zeigt sich eine signifikante Interaktion zwischen Zeit und Gruppe, $F(3,217) = 2.876$, $p < .05$, partielles $\eta^2 = .040$. Im Posttest unterscheiden sich jeweils die Value Affirmation Experimentalgruppe von der Kontrollgruppe (0.35, $p = .036$) und die Growth Mind Set Experimentalgruppe von der Kontrollgruppe (0.38, $p = .043$). Es zeigte sich ebenfalls ein signifikanter Haupteffekt der Zeit, $F(1,217) = 7{,}837$, $p = .006$, partielles $\eta^2 = .035$. Den stärksten (signifikanten) Effekt hatte die Growth Mind Set Intervention auf die Selbstwirksamkeitserwartung, $F(1,459) = 8.169$, $p = .006$, partielles $\eta^2 = .15$. Insgesamt zeigen die Ergebnisse, dass während die Kontrollgruppe einen leichten Abfall der Selbstwirksamkeitserwartung zwischen Pre- und Posttest erfuhr, sich bei den drei Experimentalgruppen ein leichter Anstieg erkennen ließ, wobei die Growth Mind Set Intervention, welche den Schüler*innen die Möglichkeit der Adaption einer dynamischen, wachstumsorientierten Grundhaltung ermöglicht, hier den nachweislich deutlichsten Effekt hatte. Die Ergebnisse werden in die bislang heterogene US-amerikanische bzw. noch wenig existente deutschen Befundlage eingeordnet. Die weiteren Analyseperspektiven der Daten, die Implikationen für zukünftige Implementationen der Interventionen in Schulen sowie die Bedeutung der Einführung dieser bislang wenig bekannten Interventionsformen in die Lehrpersonenaus- und weiterbildung werden diskutiert.

Beyond Approach-Avoidance: Relations between Achievement Goals and Regulatory Focus based on the Self-Guide Definition – A registered report

Sophie Bossert, Marc Philipp Janson
Universität Mannheim, Germany

Achievement goal theory (Dweck & Leggett, 1988; Elliot & McGregor, 2001; Urdan et al., 1997) and regulatory focus theory (Higgins, 1987, 1997, 1998) are two prominent approaches explaining individual goal striving. According to achievementgoal theory, individuals prefer different qualities of goals in achievement situations such as developing competence (mastery goals) or demonstrating competence (performance goals). These achievement goals can be subdivided into approach/avoidance on a valence dimension. Achievement goals impact numerous individual outcomes such as learner´s achievement emotions, self-efficacy, or academic achievement with mastery (approach) goals being adaptive and performance-avoidance goals maladaptive (Huang, 2011, 2012, 2016). Further, research has shown that adaptive achievement goals can be fostered in classroom settings (Ames, 1992).

Regulatory focus theory is often connected to achievement goals and proposes two motivational systems. A promotion focus emphasizes striving to reach positive outcomes such as hopes, wishes, and aspirations (ideals), a prevention focus emphasizes striving not to fail on negative outcomes, such as duties, responsibilities, and obligations (oughts). The main distinguishing factor between

regulatory foci can be seen in the valence (approach/avoidance; referred to as reference-point definition; Summerville & Roese, 2007) or the conceptual content dimension (ideals/oughts; referred to as self-guide definition). Research has connected both constructs mainly according to the reference-point definition (Janke et al., 2016; Shin et al., 2017; Świątkowski & Dompnier, 2020). We see this as a theoretical and methodological shortcoming and propose a different theoretically meaningful association of those constructs. We propose the qualitative aspect of pursuing ideals to match mastery orientations, while focusing on oughts matches performance goals. Additionally, we review literature showing mastery goals and promotion orientations to be associated with an intrapersonal perspective, and performance goals and prevention orientations with an interpersonal perspective (Habermann et al., 2022; Pekrun et al., 2014). Based on these theoretical considerations and existing studies on correlates, we propose that performance goals (approach/avoidance) are associated with a prevention focus, and mastery goals are associated with a promotion focus. A first empirical test with cross-sectional data from N = 192 participants supports our proposition of self-guide-based associations. We observed significant correlations of both performance goals with prevention (approach: r = .16, p = .02; avoidance: r = .39, p < .01) and mastery goals with promotion, r = .29, p < .01, while the correlations of mastery goals with prevention, r = -.02, p = .73, and performance goals with promotion (approach: r = .08, p = .25; avoidance: r = -.09, p = .20), representing the reference-point perspective, remained insignificant. For those insignificant results, sensitivity analyses (two-tailed) using G*POWER (Faul et al., 2007) revealed a critical $|r|$ = .14 resulting in a probability of 90% to detect significant correlations. Results of additional structural equation models were also in favor of the self-guide-based associations between achievement goals and regulatory foci compared to the reference-point-based associations. To further consolidate these findings and overcome some limitations of this first study, we designed an experiment manipulating regulatory focus and measuring achievement goals for a registered report.

We expand research on the nomological network of both constructs with our perspective and contribute to a deeper understanding of human strategic orientations. We also provide opportunities for the development of improved goal fostering interventions in achievement situations by highlighting the importance of encouraging students to pursue wishes, hopes, and aspirations (ideals) instead of focusing on duties, responsibilities, and obligations (oughts) to consolidate mastery goal orientations. Further, the current project illustrates the process of developing a registered report, an innovative open science practice to conduct thorough research based on robust theoretical considerations.

H-03: Metakognition und Feedback

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S06 S00 B41
Chair der Sitzung: Julia Hornstein

Förderung evidenzorientierten Denkens von Lehramtsstudierenden durch Peer Feedback – Effekte von Feedbackgabe- und Feedbackintegrationsprompts

Julia Hornstein, Martin Greisel, Anna Weidenbacher, Johanna Ott, Ingo Kollar
Universität Augsburg, Deutschland

Von Lehrpersonen wird zunehmend gefordert, unterrichtliche Problemsituationen im Einklang mit bildungswissenschaftlichen Theorien und Befunden zu lösen (Renkl, 2022). Entsprechende Kompetenzen sollten bereits im Lehramtsstudium aufgebaut werden (Wilkes & Stark, 2022). Empirische Studien zeigen jedoch, dass Lehramtsstudierende Schwierigkeiten haben, evidenzorientiert zu denken, etwa wenn sie die Aufgabe haben, authentische Unterrichtsprobleme zu lösen (Kiemer & Kollar, 2021). Eine Möglichkeit zur Unterstützung evidenzorientierten Denkens über authentische Unterrichtsfälle stellt die Umsetzung von Peer Feedback dar (Panadero & Alqassab, 2019). Qualitativ hochwertiges Feedback zu geben, das Peers etwa an die Aufgabenstellung erinnert (Feed Up), ihre Leistung vor diesem Hintergrund kommentiert (Feed Back) und Vorschläge für die Revision gibt (Feed Forward; Hattie & Timperley, 2007), fällt Studierenden allerdings schwer (Alemdag & Yildirim, 2022). Gleiches gilt für den effektiven Umgang mit erhaltenem Feedback (Jurkowski, 2018), insbesondere dann, wenn mehrere Feedbacks miteinander integriert werden sollen.

Die Studie untersucht daher die Effekte von Feedbackgabe- und Feedbackintegrationsprompts auf (a) die Feedbackqualität, (b) den Erwerb konzeptuellen Wissens über die dabei genutzten Evidenzen und (c) den Erwerb von Kompetenzen zum evidenzorientierten Denken. Wir nehmen an, dass beide Arten von Prompts diesbezüglich positive Effekte erzielen.

Zur Beantwortung der Fragestellung wurde eine experimentelle Studie mit einem 2x2-faktoriellen Zwischensubjektdesign mit den unabhängigen Variablen „Feedbackgabeprompts“ und „Feedbackintegrationsprompts“ durchgeführt. Es nahmen N=254 Lehramtsstudierende (77,95% weiblich, durchschnittlich im 4. Semester) mit einem Durchschnittsalter von M=22.56 (SD=4.30) Jahren teil. Die Studierenden sollten eine problematische Unterrichtssituation analysieren. Zur Analyse wurden ihnen je ein Grundlagentext zum ICAP-Modell (Chi & Wylie, 2014) und zur Cognitive Load Theory (Sweller, 1999) zur Verfügung gestellt. Darauf folgte die Feedbackgabe auf die Problemanalysen zweier Peers. Im Anschluss sollten die erhaltenen Feedbacks integriert werden, um darauf aufbauend die Ausarbeitung zu überarbeiten. Mit Blick auf die Feedbackgabe wurde variiert, ob die Studierenden Prompts erhielten, die sie dazu anhielten, dem Peer nochmals die Aufgabe zu verdeutlichen (Feed Up), seine/ihre Problemlösung konstruktiv-kritisch zu bewerten (Feed Back) und Vorschläge für die Revision zu unterbreiten (Feed Forward) oder nicht. Bezüglich der Feedbackrezeption erhielten die Studierenden je nach Bedingung entweder Prompts, die sie explizit in der Integration der erhaltenen Feedbacks anleiteten, oder unspezifische Prompts. Nach Erhalt der Erstausarbeitungen schätzten die Studierenden deren Qualität ein (F1:ω=.79,F2:ω=.88); zudem schätzten die Rezipienten das erhaltene Feedback auf seine Qualität (F1:ω=.71,F2:ω=.82) (Brooks et al., 2019) und Angemessenheit (F1:ω=.80,F2:ω=.82) (Strijbos et al., 2021) ein. Das konzeptuelle Wissen über die in den Grundlagentexten behandelten Theorien wurde über einen Wissenstest mit

Multiple-Choice-Fragen im Prä- und Posttest gemessen. Zur Messung der Kompetenz zum evidenzorientierten Denken mussten die Studierenden beispielhafte Fallanalysen nach ihrer formalen Qualität sortieren.

Bzgl. der Feedbackqualität erbrachte eine Regression hypothesenkonform einen signifikant positiven Effekt der Feedbackgabeprompts (b=.54,p=.00) und keine signifikanten Effekte der Feedbackintegrationsprompts (b=-.05,p=.77) sowie der Interaktion (b=-.14,p=.59). Auf den Wissenserwerb zeigten sich dagegen keine signifikanten Effekte (b<-.02, p>.97). Auch für den Kompetenzerwerb zeigen sich weder Haupteffekte der Feedbackgabeprompts (b=.12,p=.35) noch der Feedbackintegrationsprompts (b=.11,p=.36). Allerdings zeigte sich hier eine signifikante Interaktion (b=-.35,p=.05): Feedbackgabeprompts erzielten nur dann einen positiven Effekt, wenn zusätzlich keine Integrationsprompts präsentiert wurden. Somit stehen die Ergebnisse in Zusammenhang mit der bisherigen Forschung, die zeigt, dass Prompts einen Lerneffekt bei den Studierenden hervorrufen können (Jurkowski, 2018; Latifi et al., 2020). Allerdings zeigte sich dies eher für die Feedbackgabeprompts, nicht aber für die Feedbackintegrationsprompts. Offensichtlich reichen Prompts für die kognitiv anspruchsvolle Aufgabe der Integration mehrerer Feedbacks nicht aus, um positive Effekte zu erzielen. Der Interaktionseffekt der Prompts beim Kompetenzerwerb verweist zudem darauf, dass nicht zu viele Aspekte des Peer Feedbackprozesses gleichzeitig gepromptet werden sollten. Hierfür könnten kognitive Belastungsphänomene (Kirschner et al., 2018) verantwortlich sein. Inwiefern dies der Fall ist, sollte in künftigen Studien untersucht werden.

Förderung metakognitiver Lernstrategien in der fünften und sechsten Jahrgangsstufe

Xenia-Lea Weber[1,2], Jens Fleischer[1], Ferdinand Stebner[2], Corinna Schuster[1], Detlev Leutner[3], Joachim Wirth[1]
[1]Ruhr-Universität Bochum, Deutschland; [2]Universität Osnabrück, Deutschland; [3]Universität Duisburg-Essen, Deutschland

Selbstreguliertes Lernen (SRL) korreliert positiv mit akademischem Erfolg und gilt als Voraussetzung für lebenslanges Lernen (Dignath et al., 2008). Zur Förderung von SRL wurden verschiedene Trainings entwickelt, in denen dem Huckepack-Theorem von Klauer (2000) folgend oftmals metakognitive und kognitive Lernstrategien kombiniert trainiert werden (z.B. Stebner et al., 2015). Verschiedene Studien konnten zeigen, dass kombinierte Trainings in Bezug auf die Qualität der Anwendung kognitiver Lernstrategien (z.B. Leopold & Leutner, 2015) und die Lernleistung im Allgemeinen (z.B. Dignath & Büttner, 2015) vorteilhafter sind als nichtkombinierte Trainings. Neben der Unterscheidung von kombinierten und nichtkombinierten Trainings lassen sich auch direkte und indirekte Trainings unterscheiden. Erstere beinhalten das explizite Training von kognitiven und metakognitiven Lernstrategien, während es bei Letzteren um die implizite Unterstützung der Anwendung von Lernstrategien in einer anregenden Lernumgebung geht. Schuster und Kollegen (2019) zeigen in ihrer Interventionsstudie im Rahmen eines kombinierten Trainings mit direkter und indirekter Förderung, dass Schülerinnen und Schüler erlernte metakognitive Strategien eigenständig auf kognitive Strategien in Lernsituationen innerhalb sowie außerhalb des Trainings anwenden können, allerdings nur in sehr geringem Ausmaß. Schuster und Kollegen (2019) geben jedoch zu bedenken, dass der von ihnen eingesetzte Multi-Strategie-Test (MST; Stebner et al., 2015) womöglich nicht hinreichend sensitiv war und deshalb empirisch zu entsprechenden Bodeneffekten (M=0.36; SD=1.01) führte.

Der MST ist ein Vignetten-Test mit offenem Antwortformat. Die offenen Antworten werden mit Hilfe eines Kategoriensystems bezüglich der Anwendung kognitiver und metakognitiver Lernstrategien analysiert. Dieses Kategoriensystem wurde stark überarbeitet, um sensitiver die Anwendung von Lernstrategien erfassen zu können. In einer Re-Analyse der Daten von Schuster und Kollegen (2019) mit Hilfe des überarbeiteten Kategoriensystems zeigte sich, dass dadurch tatsächlich eine sensitivere Erfassung metakognitiver Lernstrategien erreicht und Bodeneffekte deutlich reduziert wurden (Weber et al., 2022).

Mit dem modifizierten MST soll nun zum einen geprüft werden, ob sich die Ergebnisse zur höheren Sensitivität des Instruments an einer weiteren Stichprobe replizieren lassen. Zum anderen soll der Frage nachgegangen werden, ob durch ein kombiniertes Training, im Rahmen einer direkten und indirekten Förderung, die erworbenen metakognitiven Lernstrategien transferiert werden können, sowohl auf die Bearbeitung von Trainingsaufgaben (nahe Transferaufgaben) als auch trainingsfremde Lernaufgaben (ferne Transferaufgaben) (Sweller & Paas, 2017).

An der hiesigen quasiexperimentellen Studie nahmen 456 Schülerinnen und Schüler der fünften und sechsten Jahrgangstufe (M_{Alter}=10.5, SD=0.53; 51.1% weiblich) von insgesamt sechs Schulen teil. EG1 und EG2 erhielten über ein Schulhalbjahr ein Training zum SRL (direktes Training). In dem direkten Training wurden explizit neben der Anwendung unterschiedlicher Lernstrategien auch die Kombination metakognitiver und kognitiver Lernstrategien gefördert. EG1 erhielt neben dem direkten Training auch Deutsch-Fachunterricht, der auf die metakognitiven Trainingsinhalte abgestimmt wurde (indirektes Training). Die Kontrollgruppe erhielt ausschließlich den normalen Deutsch-Fachunterricht ohne ein Training zum SRL.

Zur Erfassung von nahem und fernem Transfer wurden drei Versionen des modifizierten MST jeweils vor dem Start des Trainings und nach der Durchführung des Trainings eingesetzt. Um die Anwendung kognitiver und metakognitiver Lernstrategien zu messen, fordert der MST Schülerinnen und Schüler dazu auf, zu beschreiben, wie sie beim Experimentieren (1), bei einer Berichtigung im Fach „Deutsch“ (2) und bei der Vorbereitung auf eine Klassenarbeit allgemein (3) vorgehen würden. Während die Aufgabe beim Experimentieren eine nahe Transferaufgabe darstellt, sind die anderen beiden Aufgaben als ferne Transferaufgaben anzusehen.

Die Ergebnisse der Varianzanalysen mit Messwiederholung innerhalb des ersten Schulhalbjahres zeigen für den nahen Transfer, dass EG1 und EG2 signifikant erfolgreicher metakognitive Lernstrategien abrufen können als die Kontrollgruppe (F(2,456)=36.95, p<.001, η_p^2=.139). Damit konnte die generelle Wirksamkeit des Trainings konform zu vorigen Studien gezeigt werden (Stebner et al., 2015). Darüber hinaus zeigt sich, dass sich der von Schuster und Kollegen berichtete Bodeneffekt deutlich reduzieren ließ (M=4.69; SD=3.85). Weitere Ergebnisse, insbesondere zu fernem Transfer werden auf der Tagung präsentiert und diskutiert.

Gegenstände metakognitiver Überwachung und Zusammenhänge mit metakognitiven Urteilen

Nick Naujoks-Schober[1], Markus Dresel[2], Marion Händel[3]
[1]Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany; [2]Universität Augsburg, Germany; [3]Hochschule Ansbach, Germany

Theoretischer Hintergrund und Fragestellungen

Akkurate metakognitive Überwachung stellt eine zentrale Fähigkeit für gelingendes selbstreguliertes Lernen im Studium dar (Winne & Hadwin, 1998), aber bereitet vielen Studierenden Schwierigkeiten (Barenberg & Dutke, 2022). Zentrale Komponenten der metakognitiven Überwachung sind (1) Überwachungsstrategien, welche zielgerichtete (kognitive) Handlungen zur Beobachtung und Bewertung von Lerngegenständen umfassen (Webber et al., 1993), sowie (2) metakognitive Urteile, die Einschätzungen der eigenen Leistung darstellen (Schraw, 2009). Diese beiden Komponenten können unterschiedliche Überwachungsgegenstände fokussieren. Überwachungsstrategien zielen u.a. auf das Verständnis im Lernprozess, die Nutzung von Lernstrategien und die Bearbeitung aktueller Aufgaben ab (Greene & Azevedo, 2007). Eine systematische Betrachtung der verschiedenen Überwachungsgegenstände der Strategien erfolgte bislang nicht (Griffin et al., 2013). Daraus resultiert die erste Forschungsfrage des Beitrags (F1): Lassen sich unterschiedliche Gegenstände metakognitiver Überwachung voneinander abgrenzen? Metakognitive Urteile fokussieren im Studium hingegen häufig die erreichte Leistung in Wissensaufgaben (Rivers, 2021). Es stellt sich deshalb die Frage (F2), wie die Überwachungsstrategien miteinander zusammenhängen und ob sich stärkere Zusammenhänge zwischen Überwachungsstrategien und metakognitiven Urteilen zeigen, je ähnlicher sich die Überwachungsgegenstände beider Komponenten sind.

Methode

Die Erfassung metakognitiver Überwachung erfolgte im Rahmen einer regulären Vorlesung. Studierende (N = 171) konnten zur Mitte des Semesters an einem Übungstest mit korrektivem Feedback teilnehmen, welcher den vorangegangenen Abschnitt der Vorlesung abbildete. Der Übungstest bestand aus 14 Wissensfragen mit jeweils vier Antwortmöglichkeiten und einer richtigen Antwort. Die Überwachung während der Vorbereitung auf den Übungstest wurde vor der Testbearbeitung hinsichtlich folgender Gegenstände erfasst: (a) Überwachungsstrategien zum Verständnis und (b) Überwachungsstrategien zum strategischen Vorgehen des Lernens. Nach der Bearbeitung des Übungstests wurden die Studierenden zum (c) Einsatz von Überwachungsstrategien während der Testbearbeitung befragt. Das Ausmaß der Überwachung der Gegenstände wurde mit je drei Aussagen auf einer sechsstufigen Zustimmungsskala erfragt (1 = *stimmt gar nicht* bis 6 = *stimmt genau*). Abschließend vergaben die Studierenden ein metakognitives Urteil zu jeder Wissensfrage des Übungstests (dichotome Ja-Nein-Einschätzung zur Korrektheit der eigenen Lösung). Eine konfirmatorische Faktorenanalyse diente zur Identifikation der Überwachungsgegenstände (F1) und Pearson-Korrelationen zwischen den erfassten Variablen dienten zur Untersuchung der Zusammenhänge zwischen den Überwachungsstrategien sowie der Zusammenhänge dieser Strategien mit den metakognitiven Urteilen (F2).

Ergebnisse und ihre Bedeutung

Die Ergebnisse zur ersten Fragestellung zeigten, dass Studierende hinsichtlich (a) des Ausmaßes der Überwachung des Verständnisses, (b) der strategischen Überwachung während der Vorbereitung auf die Leistungstests sowie (c) des Ausmaßes der eingesetzten Überwachungsstrategien während der Testbearbeitung unterschieden, $\chi^2(24) = 40.32$, CFI = .93, SRMR = .06. Die Studierenden verwendeten die meisten Überwachungsstrategien während der Bearbeitung des Übungstests, M =

4.71. Bezüglich der zweiten Forschungsfrage zeigten die Korrelationen einen starken Zusammenhang zwischen den Überwachungsgegenständen während der Vorbereitung untereinander (r = .53***), während zwischen den Gegenständen der Vorbereitung und der Testbearbeitung nur schwache Zusammenhänge vorlagen (.20** ≤ r ≤ .22**). Des Weiteren zeigten sich keine signifikanten Zusammenhänge zwischen den Überwachungsgegenständen während der Vorbereitungsphase und den metakognitiven Urteilen. Ein signifikanter Zusammenhang konnte nur zwischen der Überwachung während der Testbearbeitung und den metakognitiven Urteilen nachgewiesen werden, die sich beide unmittelbar auf die Durchführung des Übungstests bezogen, r = –.29***. Die negative Korrelation lässt sich so deuten, dass Studierende, die einen häufigeren Einsatz von Überwachungsstrategien während der Testbearbeitung berichteten, eine geringere Überschätzung der eigenen Leistung aufwiesen.

Die Studie ist eine der ersten, welche die ausgeführte Ausdifferenzierung der metakognitiven Überwachungsstrategien in einem ökologisch validen Setting prüft. Die Ergebnisse verdeutlichen die Relevanz einer situationsangemessenen Betrachtung sowie Erfassung metakognitiver Überwachung und legen eine Förderung der Nutzung von Überwachungsstrategien während der Vorbereitung auf Leistungstests nahe, da in der vorliegenden Studie während der Vorbereitung weniger Überwachungsstrategien eingesetzt wurden als während der Testbearbeitung. Zur Absicherung dieser ersten Ergebnisse sollte neben einer Replikation auch eine Übertragung in andere Lernsituationen stattfinden.

Wirkung von ressourcen- und defizitorientierten Prompts beim Textlesen auf die metakognitive Genauigkeit

Stefanie Elsner, Jörg Großschedl
Universität zu Köln, Germany

Theoretischer Hintergrund

Die metakognitive Genauigkeit („metacognitive accuracy“) bei Lernprozessen lässt sich als die Übereinstimmung von subjektiven Einschätzungen und objektiven Leistungsbeurteilungen bezeichnen. Diese Genauigkeit kann als Voraussetzung für die Aufrechterhaltung von angepasstem Lernverhalten betrachtet werden. Beispielsweise werden Lernende ihre Bemühungen ein bestimmtes Lernziel zu erreichen frühzeitig beenden, wenn sie ihren eigenen Lernerfolg überschätzen. Wir gehen davon aus, dass die metakognitive Genauigkeit von der Informationsverarbeitung während des Lernprozesses abhängt. Die Informationsverarbeitung lässt sich von Prompts leiten d.h. Fragen oder visuellen Hinweisen, die der Beobachtung des Lernprozesses dienen (siehe de Boer et al., 2018; Engelmann et al., 2021; Großschedl & Harms, 2013; Zohar & Barzilai, 2013). Die Aufmerksamkeit der Lernenden lässt sich dabei zu spezifischen Merkmalen des Lernprozesses lenken. Ressourcenorientierte Prompts („Was habe ich bereits verstanden?“) und defizitorientierte Prompts („Was habe ich noch nicht verstanden?“) sind Beispiele für generische Prompts und können dabei unterstützen, die Aufmerksamkeit auf das eigene Verständnis während des Lernprozesses zu richten. Anhand des Embedded Processing Model of Working Memory (Cowan, 1988) stellen wir die Hypothese auf, dass ressourcenorientierte Prompts zu einer erhöhten und defizitorientierte Prompts zu einer verringerten metakognitiven Genauigkeit führen. Wir gehen davon aus, dass dieser Effekt in

der verstärkten Aufmerksamkeit auf Lerninhalte durch ressourcenorientierte Prompts und in der verringerten Aufmerksamkeit auf Lerninhalte („Lückenfokus“) durch defizitorientierte Prompts begründet liegt. Das vorliegende Online-Experiment testet diese Hypothese im Rahmen des Textlesens eines Fachtextes der Biologie zum Thema „Epigenetik“.

Methodik

Es nahmen 162 Studierende im Alter von 25.18 Jahren (SD: 3.62) verschiedener Lehramtsstudiengänge im Fach Biologie an dieser Studie teil. Die Teilnehmenden wurden im Rahmen eines biologiedidaktischen Pflichtmoduls rekrutiert. Dafür erhielten sie per E-Mail einen Link zur Teilnahme an der Online-Studie. Im Rahmen des Onlinefragebogens wurden das Alter, das Geschlecht, die Schulform und das Zweitfach erhoben. Anschließend erhielten die Teilnehmenden eine kurze Einführung in die Thematik „Metakognition“ (in Anlehnung an Schraw, 1998). Anknüpfend erhielten die Teilnehmenden die Aufgabe ein Buchkapitel (1466 Wörter) zum Thema „Epigenetik“ zu Herkunft, DNA-Methylierung und Unterschiede bei identischen Genen zu lesen (Knippers, 2017). Die Teilnehmenden wurden randomisiert einer von drei Gruppen zugeordnet. Eine Gruppe erhielt den ressourcenorientierten Prompt „Was habe ich bereits verstanden?“ (n = 55), eine Gruppe erhielt den defizitorientierten Prompt „Was hab ich noch nicht verstanden?“ (n = 50) und eine Kontrollgruppe erhielt keine Prompts (= 57). Die Prompts wurden jeweils an vier Stellen des Textes präsentiert. Nach dem Textlesen wurden die Teilnehmenden gebeten anzugeben, wie häufig sie die Prompts während des Textlesens nutzten. Außerdem wurden sie aufgefordert ihr Textverständnis einzuschätzen. Anschließend bearbeiteten die Teilnehmenden einen Leistungstest zum Thema „Epigenetik“. Zur Beurteilung der metakognitiven Genauigkeit wurde die Diskrepanz zwischen der subjektiven Einschätzung des Textverständnisses und dem Ergebnis des Leistungstests berechnet. Dafür subtrahierten wir die transformierten Werte beider Skalen (0 – 100%) und verglichen die Differenz zwischen den drei Gruppen. Ein Wert von 0 indiziert Übereinstimmung. Ein positiver Wert indiziert eine Überschätzung der eigenen Leistung, ein negativer indiziert eine Unterschätzung.

Ergebnisse und Bedeutung

Wir beobachteten statistisch signifikante Unterschiede in der metakognitiven Genauigkeit zwischen den drei Gruppen anhand einer Welch ANOVA für heterogene Varianzen $F(2, 92.78) = 3.57$; $p = .032$, $\eta^2 = .044$. Games-Howell post hoc Analysen zeigen eine verringerte metakognitive Genauigkeit bei defizitorientieren Prompts (M: 28.40, SD: 18.64) im Vergleich zu ressourcenorientierten Prompts (M: 18.24, SD: 21.85; $p = .03$) mit einer mittleren Effektstärke Cohen’s $d = 0.5$. Die Ergebnisse der Studie zeigen, dass der Einsatz von Prompts die metakognitive Genauigkeit verändern kann. Defizitorientierte Prompts können zu einer Überschätzung des eigenen Verständnisses im Vergleich zu ressourcenorientierten Prompts beitragen. Wir gehen davon aus, dass dieser Unterschied auf eine Verschiebung von Aufmerksamkeit bzw. der veränderten Aufrechterhaltung von Informationen im Arbeitsgedächtnis während des Textlesens zurückführbar ist.

H-04: Herausforderungen in der Hochschulbildung

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S06 S02 B06
Chair der Sitzung: Franziska Baier-Mosch

Heterogenität selbst erleben – Replikation eines Experiments zur Überzeugungsveränderung von Lehramtsstudierenden

Franziska Baier-Mosch, Charlotte Dignath, Mareike Kunter
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Germany

Theoretischer Hintergrund & Fragestellung

Ein an die heterogenen Voraussetzungen der Schüler*innen angepasster Unterricht ist in der deutschen Unterrichtspraxis eher selten (Helmke, 2010; Warwas et al., 2011), obwohl seine Bedeutsamkeit immer wieder betont wird (Klieme & Warwas, 2011; Trautmann & Wischer, 2008). Gründe für die geringe Umsetzung können neben kontextuellen Faktoren (Helmke, 2010) lehrkraftbezogenen Faktoren wie ungünstigen Überzeugungen sein (Warwas et al., 2011). Forschung zeigt, dass positive Lehrkraft-Überzeugungen, z.B. in Hinblick auf die *Relevanz* eines Themas, bedeutsame Prädiktoren für Unterrichtshandeln sind (van Aalderen-Smeets et al., 2015). Aufgrund der Relevanz positiver Lehrkraftüberzeugungen, ist es zentral, diese zu fördern. Interventionen, die die Entwicklung positiver Überzeugungen (angehender) Lehrkräfte z.B. in Hinblick auf Beeinträchtigungen von Schüler*innen (z.B. Dyslexie) adressieren, umfassen das eigene Erleben dieser Beeinträchtigungen (Wadlington et al., 2008). Als Wirkmechanismus wird dabei Empathie angenommen. Die aktuelle Studie überträgt eine solche Intervention auf das eigene Erleben von Lehramtsstudierenden, Teil einer heterogenen Lerngruppe zu sein, um deren Relevanzüberzeugungen in Hinblick auf die Berücksichtigung von Schüler*innenheterogenität im späteren eigenen Unterricht (AV) zu fördern.

Methode

In einer ersten experimentellen Studie (Studie 1) wird untersucht, ob eine Intervention, in der Lehramtsstudierende sich selbst als Teil einer heterogenen Studierendengruppe erleben (Studierende lesen Text zum Thema Lerntheorien, beurteilen die Schwierigkeit des Texts und lesen danach simuliertes *heterogenes* Feedback zur Text-Schwierigkeit von vermeintlichen Kommiliton*innen; EG1) effektiver ist als eine Intervention, in der ein Lehrbuchtext die bestehende Heterogenität von (Lehramts-)Studierenden lediglich beschreibt (EG2) und eine Intervention, die sich nicht mit dem Thema Heterogenität von Lehramtsstudierenden beschäftigt (Studierende lesen lediglich Text über Lerntheorien; KG). In einer zweiten experimentellen Studie (Studie 2; Replikationsstudie) wird untersucht, ob sich die Befunde aus Studie 1 anhand einer neuen Stichprobe mit Lehramtsstudierenden replizieren lassen.

An Studie 1 nahmen 156 Lehramtsstudierende aller Schulformen (77% weiblich) teil. Als AVs (prä-post) wurden die Relevanzüberzeugungen der Lehramtsstudierenden in Hinblick auf die Berücksichtigung von Schüler*innenheterogenität im Unterricht (5 Items; $\alpha_{prä/post}$ = .69/.81) und die Berücksichtigung von Studierendenheterogenität in der Hochschule (4 Items; $\alpha_{prä/post}$ = .82/.85) mit Hilfe von Fragebögen erfasst. An Studie 2 (Replikationsstudie) nahmen 152 Lehramtsstudierende aller Schulformen (78% weiblich) teil und es wurden ebenfalls die jeweiligen Relevanzüberzeugungen anhand der gleichen 5 Items ($\alpha_{prä/post}$ = .62/.71) bzw. 4 Items (; $\alpha_{prä/post}$ = .73/.88) erfasst.

Ergebnisse & ihre Bedeutung

In Studie 1 zeigten sich positive Veränderung in den Relevanzüberzeugungen zur Berücksichtigung von Studierendenheterogenität in der EG1 (η^2_p =.28) und EG2 (η^2_p =.12) sowie einen positive Veränderung in den Relevanzüberzeugungen zur Berücksichtigung der Schüler*innenheterogenität in der EG1 (η^2_p =.21). In Studie 2 konnten die Ergebnisse weitestgehend repliziert werden, die Effekte fielen jedoch kleiner aus (Effektstärke η^2_p). Die positive Veränderung in den Relevanzüberzeugungen zur Berücksichtigung von Studierendenheterogenität lag in der EG1 bei η^2_p =.16 und in der EG2 bei η^2_p =.11. Die positive Veränderung in den Relevanzüberzeugungen zur Berücksichtigung der Schüler*innenheterogenität lag in der EG1 bei η^2_p =.07.

Die Arbeit zeigt, dass eine kurze Intervention, die das eigene Erleben von Lehramtsstudierenden, Teil einer heterogenen Lerngruppe zu sein, nutzt, positive Effekte auf die Relevanz, die Lehramtsstudierende der Berücksichtigung von Schüler*innenheterogenität im eigenen Unterricht beimessen, haben kann. Der Effekt ist allerdings eher klein und in einer Replikationsstudie fällt er kleiner aus als in einer ersten Studie. Dies zeigt, dass weitere Replikationsstudien notwendig sind, um den Effekt weiter empirisch abzusichern und verdeutlicht somit die Mühe, die es erfordert, wenn vorläufiges Wissen in gut abgesichertes Wissen überführt werden soll (siehe Tagungsmotto).

Es sind weitere Studien notwendig, die untersuchen, ob und inwiefern die höhere wahrgenommene Themenrelevanz durch die Lehramtsstudierenden anschließend genutzt werden kann, um eine tiefere Auseinandersetzung der Studierenden mit dem entsprechendem Lernstoff (Heterogenität von Schüler*innen und adaptiver Unterricht) in der Lehrerbildung anzuregen.

An extended Theory of Planned Behaviour to predict university students intentions towards social exclusion

Mareike Brehmer[1], Jennifer Meyer[2]
[1]University of Agder, Norway; [2]Leibniz Institute for Science and Mathematics Education (IPN), Germany

Introduction

Social exclusion from group activities is a prevalent issue among students in higher education and poses challenges to the creation of a safe and inclusive learning environment. We conceptually replicated and extended a study by Hayashi and Tahmasbi (2021) to investigate to what extent the Theory of Planned Behaviour may predict university students' behavioural intention towards in-person peer social exclusion in a hypothetical group work setting.

Relevance

Social exclusion describes an individual purposefully being left out of a group setting (Williams & Govan, 2013). The ways onlooking people react to social exclusion and other bullying incidents are various (e. g. Salmivalli, 2014; Bauman et al., 2020; Harrison et al., 2022). Assertive onlookers intervene in a situation in a prosocial way and can influence group norms (Abbott & Cameron, 2014). The present study will enhance knowledge about factors predicting prosocial and antisocial behavioural intentions to intervene in a hypothetical situation with the aim to support resource-oriented pedagogical work in higher education (Lambe & Craig, 2020). Fostering factors which contribute to prosocial behavioural intentions in students may contribute to their acting as assertive

onlookers who challenge antisocial behaviours among their fellow students and encourage a safe social environment from within the peer group.

Theoretical background

The Theory of Planned Behaviour assumes that a person's attitudes towards a behaviour, their subjective norms deriving from significant others, and their perceived behavioural control predict their intention to execute the behaviour in question (Ajzen, 2012). The intention is considered a proxy and the most reliable predictor of behaviour (Fishbein & Ajzen, 2010). The Theory of Planned Behaviour is open to extensions and several potential predictors have been examined previously. In the above mentioned study by Hayashi and Tahmasbi (2021) empathy and anticipated regret were found to be significant predictors of intention, over and above the standard variables.

Hypotheses

In line with their conceptualisation, we included these in our replication study and hypothesised the following for preliminary analyses:

$H_{1a\text{-}1c}$: The more positive the attitude, subjective norms, and perceived behavioural control, the likelier a person intends to intervene

H_2: Higher empathic concern will significantly predict intention over and above attitude, subjective norms, and perceived behavioural control

H_3: Higher anticipated regret will significantly predict intention over and above attitude, subjective norms, and perceived behavioural control.

Methods and sample

We collected data using an online questionnaire distributed to undergraduate students. The final sample consisted of N=420 undergraduate students (M_{Age}= 22.76 years, SD_{Age}= 4.02 years; 47.0% male, 50.6 % female, 2.4 % other). Following a text describing a hypothetical group work situation asking participants to imagine being an onlooker of social exclusion, the intention to intervene (4 items, α =.90), attitude (7 items, α =.77), subjective norms (6 items, α =.77), perceived behavioural control (5 items, α=.76), Davis' (1980) empathic concern (7 items, α=.80) and Hayashi and Tahmasbi 's (2021) anticipated regret (2 items, r=.74**) were assessed.

Preliminary analyses and results

Preliminary regression analyses were conducted in IBM SPSS 28. The data supported $H_{1a\text{-}1c}$, indicating that attitude (*ß*=.312), subjective norms (*ß*=.220) and perceived behavioural control (*ß*=.252) were significant (all *p*<.001) predictors of intention. However, entering empathic concern and anticipated regret led to a strong decrease of the prediction by subjective norms (*ß*=.084; *p*=.057) which became insignificant. While anticipated regret (*ß*=.307; p<.001) significantly predicted intention (as H_3 assumed), empathic concern did not (*ß*=.056; *p*=185) (contrary to H_2).

Outlook towards final paper

Subsequent more complex analyses based on latent constructs will further examine the relationships among the standard and additional predictor variables and consider more specified outcome variables

describing prosocial and antisocial behavioural intentions towards incidents of social exclusion in university students.

Ziele in der Promotionsphase: Exploration ihrer Entwicklung und Zusammenhänge mit Kontextmerkmalen

Martin Daumiller[1], Charlotte Seidel[1], Stefan Janke[2], Oliver Dickhäuser[2], Markus Dresel[1]
[1]Universität Augsburg, Germany; [2]Universität Mannheim, Germany

Theoretischer Hintergrund

Wissenschaftler*innen haben die gesellschaftlich bedeutende Funktion, kontinuierlich neue Erkenntnisse zu gewinnen. Wesentlich ist hierfür ihre berufliche Motivation, die durch ihre Zielorientierungen charakterisiert werden kann (Daumiller et al., 2020). Erkenntnisse der Zielorientierungsforschung sind allgemein dadurch limitiert, dass häufig ein reduziertes Verständnis von Zielorientierungen als statische Personenmerkmale zugrundeliegt. Es gibt deutliche Erkenntnisdefizite zu ihrer Entwicklung jenseits des Kindes- und Jugendalters und deren Bedingungsfaktoren.

Inhaltlich lassen sich Masterziele und Performanzziele unterscheiden. Masterziele umfassen Aufgabenbewältigung (*Aufgabenziele*) und persönliche Kompetenzverbesserung (*Lernziele*). Performanzziele beziehen sich auf die Demonstration von Kompetenz insofern sie auf das Erreichen erwünschter (*Annäherungsziele*) oder das Vermeiden unerwünschter Zielzustände (*Vermeidungsziele*) ausgerichtet sind oder sich auf das Übertreffen anderer (*Normziele*) oder die Erscheinung vor anderen (*Erscheinungsziele*) beziehen.

Wesentliche Veränderungen dieser Ziele sollten auch im Erwachsenenalter und dabei insbesondere in markanten Transitionsphasen auftreten (Becker et al., 2017). Wissenschaftler*innen sind dahingehend wenig beforscht, eignen sie sich aber besonders, um solche Veränderungen zu untersuchen, denn markante berufliche Transitionen lassen sich hier klar und normativ benennen. Eine wesentliche Phase ist die Promotionsphase. Erwartbar ist für diese im Sinne motivationaler Kosten des Berufseinstieges eine Abnahme von Masterzielen, da bestehende Sachzwänge des Arbeitsumfeldes nach der Transition in den Beruf unmittelbar erlebt werden und in Folge die Lernfreude abfällt (Klassen et al., 2014).

Promotionen laufen allerdings nicht gleich ab, sondern sind durch unterschiedliche Rahmenbedingungen gekennzeichnet, die auch für die motivationalen Entwicklungsprozesse bedeutsam sein dürften. Das lässt interindividuelle Unterschiede in der Zielentwicklung in der Promotionsphase erwarten, wobei sich Zielorientierungen unterschiedlich entwickeln sollten, je nachdem wie sehr Aspekte des Arbeitskontexts das Lernen oder Konkurrenz bzw. Wettbewerbsdruck betonen (vgl. Meece et al., 2006).

Fragestellung

Wir untersuchten allgemeine Entwicklungstrends der Zielorientierungen in der Promotionsphase und explorierten inwieweit sich unterschiedliche Entwicklungsverläufe auf unterschiedliche Rahmenbedingungen zurückführen lassen. Hierzu fokussierten wir auf klar benennbare

Kontextmerkmale, die relevante Impulse für die Zielentwicklung setzen könnten (Zielvereinbarungen, Gratifikationen, Verpflichtungen zum Austausch oder Weiterbilden).

Methode

Datengrundlage waren 137 Promovierende (davon 79 weiblich; durchschnittliches Alter: *M*=29.4, *SD*=5.0, Jahre) unterschiedlicher Fachbereiche. Befragte machten mittels eines incentivierten Paper-and-Pencil Fragebogens viermal, alle 6 Monate, Angaben zu ihren Zielen und zu Messzeitpunkt 3+4 zu den Kontextmerkmalen.

Mit je 4 Items von Daumiller und Dresel (2020) erfassten wir *Annäherungsaufgabenziele* und *Annäherungslernziele* sowie *Annäherungserscheinungs-*, *Vermeidungserscheinungs-*, *Annäherungsnorm-* und *Vermeidungsnormziele* (gute interne Konsistenz: ω_h=.83–.92).

Hinsichtlich des relevanten Arbeitskontexts (z.B. Arbeitsgruppe, Fachbereich) erfassten wir das Vorliegen von *Zielvereinbarungen* zu Anzahl und Art erwarteter Publikationen, *Publikmachungen* der Forschungsleistung (z.B. hinsichtlich Publikationen, Drittmittel), an Forschungsleistung geknüpfte *Gratifikationen* (z.B. Preise, Gehalt), *verpflichtenden Austausch* mit Kolleg*innen und *verpflichtende Teilnahme an methodischen Weiterbildungen (*sehr gute Retest-Reliabilität: *r*=.91-95).

Wir berechneten individuelle lineare Entwicklungstrends mittels latenter Wachstumskurvenmodellierung und regressierten diese auf die erfassten Kontextmerkmale in einem Strukturgleichungsmodell. Quadratische Entwicklungstrends lieferten keine signifikant bessere Beschreibung der Entwicklungsverläufe.

Ergebnisse und Bedeutung

Grundlegend zeigten die Ergebnisse erwartungskonform eine signifikante Abnahme der Lernziele über die Promotionsphase (*t*=-2.12, *p*=.03). Die linearen Entwicklungstrends der anderen Ziele waren nicht signifikant von Null verschieden, hatten jedoch hohe intra-individuelle Varianz (*Var*=1.3-1.8, *p*<.01). Dies bedeutet, dass unterschiedliche Entwicklungen für die Promovierenden vorlagen.

Diese verschiedenen Veränderungen hingen wiederum systematisch mit den betrachteten Kontextaspekten zusammen. So entwickelten sich bei Promovierenden ohne Zielvereinbarungen Lernziele ungünstiger als bei Promovierenden mit Zielvereinbarungen; Publikmachungen der Forschungsleistung waren mit einer vermehrten Zunahme von Annäherungsnormzielen assoziiert, Gratifikationen bei guter Forschungsleistung gingen indes mit einer Reduktion von Annäherungserscheinungszielen einher. Verpflichtender Austausch mit Kolleg*innen war mit einer Reduktion von Vermeidungsnormzielen und Annäherungsaufgabenzielen assoziiert, verpflichtende Weiterbildungsveranstaltungen mit einer Zunahme an Aufgabenzielen.

Diese Befunde bestärken, dass in der Promotionsphase relevante Veränderungen von Zielorientierungen vorliegen, wobei diese durch individuelle Entwicklungstrends gekennzeichnet sind und systematisch mit unterschiedlichen Kontextaspekten korrespondieren. Dies verweist auf die Relevanz kritischer Transitionsphasen für die Untersuchung von Kontexteffekten im Rahmen eines umfassenderen Verständnisses der Zielentwicklung über die Lebensspanne.

Profile epistemologischer Überzeugungen, deklarativen Wissens über wissenschaftliche Erklärungen und Erklärungskompetenz und deren Veränderung nach einer Intervention – Ergebnisse einer latenten Profiltransitionsanalyse

Eric Klopp, Theresa Krause-Wichmann, Robin Stark
Universität des Saarlandes, Deutschland

Theoretischer Hintergrund

Wissenschaftliche Erklärungen führen beobachtete Sachverhalte auf wissenschaftliche Theorien zurück. Die Kompetenz, wissenschaftliche Erklärungen zu generieren (Erklärungskompetenz; Klopp & Stark, 2018), ist ein wesentliches Element wissenschaftlicher Kompetenz (Dietrich et al., 2015). Da der strukturelle Aufbau einer wissenschaftlichen Erklärung strukturgleich mit der Konstruktion eines Arguments ist, stellen epistemologische Überzeugungen wichtige Einflussfaktoren von Erklärungskompetenz bzw. von deklarativem Wissen über die Struktur von Erklärungen und für diese wichtigen Normen dar (deklaratives Erklärungswissen; Klopp et al., 2016; Kuhn, 1991). Beispielsweise können Überzeugungen zur Rechtfertigung durch Autoritäten die Wahl einer für eine Erklärung zutreffenden oder nichtzutreffenden Theorie bedingen. Starke Überzeugungen zur Rechtfertigung durch multiple Quellen können der Abgrenzung der zur Erklärung benutzten Theorie gegenüber Alternativtheorien zuträglich sein. Aus einer personenzentrierten Sichtweise (Bergman et al., 2003) stellt sich die Frage, ob es Profile epistemologischer Überzeugungen gibt, die mit hoher Erklärungskompetenz und ausgeprägtem deklarativem Erklärungswissen einhergehen. Erklärungskompetenz kann gefördert werden, z.B. durch Lernen aus advokatorischen Fehlern (Wagner et al., 2014). Epistemologische Überzeugungen können auch für Lernprozesse relevant sein, z.B. können starke Überzeugung zur Sicherheit des Wissens dem Lernprozess abträglich bzw. starke Überzeugungen zur reflexiven Natur von Wissen den Lernprozess unterstützen. Somit stellt sich die Frage, wie sich die Profile epistemologischer Überzeugungen und Erklärungskompetenz nach einer Förderintervention verändern.

Fragestellung

Die Studie untersucht, ob sich 1) aus personenzentrierter Perspektive verschiedene Profile epistemologischer Überzeugungen, deklarativen Erklärungswissens und Erklärungskompetenz ergeben und 2) wie sich diese Profile nach einer Intervention zur Förderung deklarativen Erklärungswissens und Erklärungskompetenz ändern.

Methoden

Die Stichprobe bestand aus 108 Studierenden, die eine auf dem Lernen aus advokatorischen Fehlern basierende Lerneinheit zur Förderung von Erklärungskompetenz (EK) und deklarativem Erklärungswissen (DEW) erhielten. In einem Prä-Posttest-Design wurde sechs Dimensionen epistemologischer Überzeugungen (eÜ) erfasst: Persönliche Rechtfertigung (PR; αPrä=.66, αPost=.76), Rechtfertigung durch Autorität (RA; αPrä=.81, αPost=.79), Rechtfertigung durch multiple Quellen (RQ; αPrä=.81, αPost=.82), Rechtfertigung durch die Scientifc community (RC; αPrä=.62, αPost=.75), Sicherheit des Wissens (SW; αPrä=.70, αPost=.72), und Reflexive Natur des Wissens (RN; αPrä=.62, αPost=.75). Mittels eines Multiple-Choice-Tests wurde DEW erfasst (αPrä=.89 αPost=.73). EK wurde mittels einer offenen Aufgabe erfasst (Klopp & Stark, 2018).

Ergebnisse

Eine latente Profiltransitionsanalyse ergab eine 2-Profil-Lösung sowohl für den ersten als auch den

zweiten Messzeitpunkt. Die beiden Profile zum ersten Messzeitpunkt lassen sich als gute Erklärer mit günstigen eÜ (GE/GEÜ) sowie als schlechte Erklärer mit ungünstigen eÜ (SE/UEÜ) interpretieren und unterscheiden sich signifikant in EK. Zudem zeigen die GE/GEÜ eine signifikant geringere Ausprägung von RA und eine höhere Ausprägung von RQ. Die Profile zum zweiten Messzeitpunkt lassen ich als gute Erklärer (GE) und sehr gute Erklärer (SGE) interpretieren, die sich signifikant sowohl in ihrer EK als auch ihrem DEW unterscheiden, nicht aber in ihren eÜ. Beim Übergang von GE/GEÜ nach GE erfolgt ein signifikanter Zuwachs an DEW und ein signifikanter Rückgang in PR. Beim Übergang von GE/GEÜ nach SGE zeigt sich eine signifikante Erhöhung an EK und DEW aber keine Veränderung von eÜ. Beim Übergang von SE/UEÜ nach GE zeigen sich signifikante Zuwächse an EK und DEW sowie in RQ und RN und überdies ein signifikanter Rückgang in RA. In gleicher Weise zeigen sich beim Übergang von SE/UEÜ nach SGE signifikante Zuwächse an EK und DEW sowie in RQ und SW sowie ein signifikanter Rückgang in SW.

Zum ersten Messzeitpunkt lässt sich somit ein Muster epistemologischer Überzeugungen aufzeigen, das günstig für wissenschaftlicher Erklärungen ist. Zudem erfolgen Übergänge zwischen den Profilen, die konform mit den Überlegungen zur Relevanz epistemologischer Überzeugungen für Erklärungskompetenz sind. Ein überraschendes Ergebnis hingegen ist, dass sich die beiden Profile am zweiten Messzeitpunkt nicht in den epistemologischen Überzeugungen unterscheiden. Eine Interpretation könnte sein, dass die Intervention epistemologische Überzeugungen dergestalt ausgleicht, dass diese in beiden Profilen dem für Erklärungen günstigen wissenschaftlichen Denken zuträglich sind.

H-05: Evidenz, Klausurnoten und Ausbildungserfahrung

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S06 S03 B06
Chair der Sitzung: Jürgen Schneider

Vermittlung von Effektgrößen an Lehrkräfte: Untersuchung verschiedener Visualisierungen und ihrer Anreicherungsoptionen zur Förderung des Verständnisses

Jürgen Schneider[1], Kirstin Schmidt[2], Kristina Bohrer[2], Samuel Merk[2], Iris Backfisch[3], Salome Wagner[3]
[1]DIPF, Germany; [2]University of Education Karlsruhe; [3]University of Tübingen

Aus der Perspektive einer evidence-informed practice ist das Wissen darüber, "what works", eine Schlüsselinformation für Lehrpersonen (Brown et al., 2017). Was in der Bildungsforschung funktioniert, wird in der Regel durch *treatment efficacy* gemessen (Slavin, 2020). Daher sind die Effektgrößen aus synthetisierter Forschung eine Schlüsselinformation, die zwischen Forschung und Lehrpersonen vermittelt werden muss. Sogenannte *Clearing Houses* verwenden meist textuelle Darstellungen von Effektgrößen, obwohl visuell dargestellte statistische Informationen unter bestimmten Bedingungen möglicherweise akkurater und schneller abgerufen werden können (Franconeri, 2021). Diesbezüglich gibt es erste empirische Hinweise darauf, dass die akkurate Wahrnehmung von Informationen (z. B. Streuung), je nach *Art der verwendeten Visualisierung* variiert (Pierce & Chick, 2013). *Anreicherungsoptionen* für diese Visualisierungen können die komplexen Prozesse der Auseinandersetzung mit Effektgrößen weiter unterstützen. Ein gängiger Ansatz besteht darin, standardisierte Benchmarks bereitzustellen, an denen Forschende ihre Interpretationen typischerweise ausrichten (Cohen, 1988). Darüber hinaus legt Forschung zum multimedialen Lernen nahe, dass die Verwendung von cues die Aufmerksamkeit der Betrachtenden lenken kann. Dieser sogenannte signaling-effect kann das Verständnis erhöhen und die kognitive Belastung verringern (Schneider et al., 2018). Bislang ist noch unklar, inwieweit Visualisierungen von Effektgrößen die Wahrnehmung, sowie das Verständnis von Effektgrößen und die Zuschreibung von praktischen Implikationen beeinflussen.

Mit einer ersten explorativen Studie werden die folgenden Forschungsfragen beantwortet:

Beeinflusst die Art der Visualisierung die Wahrnehmung, das Verständnis und die praktischen Implikationen bei der Bewertung von Effektgrößenvisualisierungen durch Lehrpersonen? (RQ 1a)
Welche Art der Visualisierung führt dabei zu den wünschenswertesten Einschätzungen? (RQ 1b)

Auf der Grundlage der Ergebnisse dieser explorativen Studie werden wir einen registrierten Bericht mit den folgenden Forschungsfragen durchführen:

Wie wirken sich Effektgrößenvisualisierungen mit (vs. ohne) Benchmarks auf die Wahrnehmung, das Verständnis und die praktischen Implikationen aus? (RQ 2a)

Wie wirken sich Effektgrößenvisualisierungen mit (vs. ohne) textliche Hinweise auf die Wahrnehmung, das Verständnis und die praktischen Implikationen aus? (RQ 2b)

Methode

Delphi-Studie (abgeschlossen)

Um die große Anzahl möglicher Visualisierungen einzugrenzen, haben wir eine Delphi-Studie mit Experten aus den Bereichen Datenvisualisierung (n1 = 4) und Wissenschaftskommunikation in Clearinghäusern (n2 = 4) durchgeführt.

Studie 1 (laufend)

Für unsere explorative Studie 1 befragen wir aktive Lehrpersonen über das Online Survey Tool Prolific. Um die Stichprobengröße zu bestimmen, nehmen wir Bayes'sches Updating vor, bis wir ein zufriedenstellendes Evidenzniveau erreichen. Dieses setzen wir a priori auf BF > 5 oder BF < 1/5.

Um RQ 1a und RQ 1b zu untersuchen, werden wir ein 4 x 6 within-person-experiment mit den beiden experimentellen Faktoren Visualisierungstypen (siehe Ergebnisse Delphi-Studie) und sechs Effektgrößen durchführen.

Wir testen derzeit etablierte Items (z.B. wahrgenommene Informativität; Lortie-Forgues et al., 2021) und neue Operationalisierungen (Sensitivität, Akkuratesse), um zu entscheiden, welche Messungen präregistriert werden.

Studie 2 (bevorstehend)

In der konfirmatorischen Studie 2 werden wir die Auswirkungen der beiden Anreicherungsoptionen Benchmarks und textuelle Hinweise unter Verwendung des "erfolgreichsten" Visualisierungstyps aus Studie 1 untersuchen. Die Stichprobenakquise und eingesetzte Instrumente gleichen Studie 1.

Um RQ 2a und RQ 2b zu untersuchen, werden wir ein experimentelles Design mit den beiden Faktoren Benchmarks (mit Benchmarks vs. ohne Benchmarks) und textuellen Hinweisen (mit Hinweisen, die Überschneidungen signalisieren vs. mit Hinweisen, die Unterschiede signalisieren vs. ohne Hinweise) verwenden.

Ergebnisse

In der Delphi-Studie identifizierten beide Expertengruppen den Raincloud Plot (Gruppen auf der y-Achse), den Gardner Altman Plot (Gruppen auf der x-Achse) und Halfeye Plots (Gruppen auf der y-Achse und der x-Achse) als die am besten geeigneten Visualisierungstypen, um den Lehrkräften Effektgrößen zu vermitteln. Diese Visualisierungen werden in unserer explorativen Studie 1 verwendet, um Frage 1a und Frage 1b zu beantworten. Die Ergebnisse aus Studie 1 werden in der konfirmatorischen Studie 2 verwendet. Weitere Ergebnisse werden auf der Konferenz vorgestellt.

Zwei Studien, ein Ergebnis: Lehramtsstudierende unterliegen im Umgang mit Evidenz dem Ankereffekt

Kristina Bohrer, Kirstin Schmidt, Samuel Merk
Pädagogische Hochschule Karlsruhe, Germany

Einleitung

Lehrpersonen sind dazu angehalten, in ihrer professionellen Praxis mit verschiedenen Arten von Evidenz umzugehen, die von Ergebnissen der Bildungsforschung bis hin zu Primärdaten wie lokale Schuldaten reichen (Brown et al., 2017; Schildkamp, 2019). Dies wird von verschiedenen Barrieren erschwert, zu denen auch kognitive Verzerrungen wie Ankereffekte zählen. Der Ankereffekt tritt dann auf, wenn ein numerisches Urteil an einen bereits bekannten Vergleichsmaßstab ausgerichtet wird (Röseler & Schütz, 2022; Tversky & Kahneman, 1974). Dieser Effekt wurde bisher im Rahmen evidenzinformierter Lehrpraxis kaum untersucht, weshalb zwei experimentelle Studien mit Lehramtsstudierenden durchgeführt wurden.

Studie 1

In der ersten Studie wurde analysiert, inwieweit verschiedene Anker die *Evaluation von Forschungsergebnissen* beeinflussen. Hierfür wurden 233 Lehramtsstudierende (85% weiblich, durchschnittlich im 3,36 Semester) in einem randomisierten kontrollierten Experiment sequentiell und randomisiert zwei kurze Forschungsberichte zu bildungswissenschaftlichen Themen präsentiert (Lernstile, Entdeckendes Lernen). Innerhalb dieser wurde unter anderem über die gezogene Stichprobengröße informiert (randomisiert $N = 15$ oder $N = 500$), über die die Teilnehmenden anschließend urteilen sollten, ob sie zur Beantwortung der Forschungsfrage angemessen ist (Likert Skala von 1 = *stimme gar nicht zu* bis 7 = *stimme voll und ganz zu*).

Dieses Vorgehen resultierte in insgesamt vier Gruppen: Zwei Kontrollgruppen, die in beiden Berichten dieselbe Stichprobengröße erhielten ($N = 15$ oder $N = 500$) und zwei Experimentalgruppen, die randomisiert zwei unterschiedliche Stichprobengrößen erhielten ($N_1 = 15$, $N_2 = 500$ oder umgekehrt). Dabei wurde angenommen, dass die Lehramtsstudierenden die Angemessenheit der zweiten und größeren Stichprobengröße ($N_2 = 500$) als höher bewerten, wenn sie zuvor den kleinen Anker erhalten haben ($N_1 = 15$). Und umgekehrt wurde angenommen, dass der große Anker ($N_1 = 500$) zu einer geringeren Bewertung der Angemessenheit der zweiten Stichprobengröße ($N_2 = 15$) führt, während in den Kontrollgruppen kein Unterschied zu erwarten ist. Eine Bayesianische ANOVA zeigt starke Evidenz für große Ankereffekte (d = 2.97 , BF10 > 100).

Studie 2

In der zweiten experimentellen Studie wurde untersucht, inwieweit Ankereffekte das *Erheben von Daten* durch Lehramtsstudierende beeinflussen. Hierfür durchliefen 38 Lehramtsstudierende (70,59% weiblich; durchschnittlich im 4,73 Semester) drei virtuelle Studien, um unterschiedliche Forschungsfragen zu beantworten. Die erste virtuelle Studie war ein fiktives Übungsbeispiel, während sich die anderen beiden auf unterschiedliche pädagogische Themen bezogen (z.B. Benotung von Jungen und Mädchen). Der Aufbau jeder Studie war jedoch identisch: Zunächst wurde das Thema vorgestellt (einschließlich einer Forschungsfrage, z. B. "Erhalten Mädchen bessere Noten als Jungen?") und dann das Studiendesign erläutert (z. B. ein Journalist, der nach dem Zufallsprinzip Telefoninterviews durchführte). Anschließend zogen die Teilnehmenden sequenziell Stichproben, um die Forschungsfrage zu beantworten. Nach jeder Ziehung konnten die Lehramtsstudierenden auf der Basis einer Grafik entscheiden, ob sie zur Beantwortung der Forschungsfrage bereits genug Daten gesammelt hatten. Insgesamt konnten sie so lange Stichproben ziehen, bis sie sich sicher waren, eine adäquate Antwort auf die Forschungsfrage geben zu können.

Um Ankereffekte messen zu können, wurde im zweiten virtuellen Experiment der Modus des Stichprobenziehens manipuliert (unabhängige Variable). Die Teilnehmenden erhielten vordefinierte unterschiedliche Antwortoptionen (Gruppe 1: $N = 4$, $N = 8$ oder $N = 16$ [kleiner Anker] vs. Gruppe 2: ($N = 5$, $N = 50$, $N = 100$ und N = 200 [kleiner bis großer Anker]). Im dritten Experiment konnten sie den Stichprobenumfang erneut frei festlegen (abhängige Variable).

Bayesianische Wilcoxon-Tests zeigen starke Evidenz für große Ankereffekte (d = 1.26, BF10 > 100).

Diskussion

Beide Studien zeigen Evidenz für Ankereffekte, wenn Lehramtsstudierende mit verschiedenen Formen von Evidenz umgehen. Die Relevanz dieser Ergebnisse lässt sich in einem kurzen Beispiel skizzieren: Wenn Lehrpersonen und Lehramtsstudierende überwiegend Studien mit großen Stichproben kennen, kann das als ein Anker wirken und zu einer Abwertung von (z.B. experimentellen) Studien mit kleineren Stichprobengrößen führen obwohl diese ausreichend statistische Power aufweisen.

Vorhersage von Klausurnoten im Mathematikstudium mit Machine Learning

Martin Fröhlich, Stefan Krauss, Sven Hilbert
Universität Regensburg, Deutschland

Die hohe Abbrecherquote bei Studiengängen mit Mathematik legt eine Untersuchung der Bedingungen nahe, die einen Erfolg im Mathematikstudium begünstigen (Heublein, 2020). Eine valide Messgröße für den Erfolg im Mathematikstudium liefern die erreichten Noten in Klausuren. Um in einer Klausur erfolgreich zu sein, sind begleitend zum Besuch einer Vorlesung sowohl kontinuierliche Lernstrategien während des Semesters sowie spezifisch für die Klausurvorbereitung als auch adäquates Kooperationsverhalten mit anderen Studierenden maßgeblich entscheidend.

Im Vortrag wird die Untersuchung der folgenden beiden Forschungsfragen vorgestellt: (1) Kann mithilfe von selbstberichteten Lernstrategien und Kooperationsverhalten die Klausurnote von Studierenden vorhergesagt werden? (2) Welche Lernstrategien und welches Kooperationsverhalten führen dabei zu den besten Klausurnoten?

An der Universität Regensburg wurden dafür im Sommersemester 2021 und im Sommersemester 2022 jeweils die Studierenden der Vorlesung Lineare Algebra II bei demselben Dozenten im Hinblick auf deren Lernstrategien und ihr Kooperationsverhalten mit Kommiliton*innen befragt. Insgesamt wurden hierzu 150 Selbsteinschätzungen mittels eines Fragebogens, der aus einer informellen Delphi-Studie innerhalb der Mathematikfakultät sowie ausgewählten Items aus dem *LimSt* (Liebendörfer et al., 2021) konstruiert wurde, zu drei Messzeitpunkten erhoben. Zusätzlich wurde ein Prätest zum Vorwissen zu Semesterbeginn und die Punktzahl in der Klausur am Semesterende als Posttest erfasst. 96 Studierende wurden im Sommersemester 2021 und 83 Studierende im Sommersemester 2022 befragt.

Zur Beantwortung der beiden Forschungsfragen wurden – statt der üblichen Skalenbildung – mittels Machine Learning Verfahren *alle* 150 Items *gleichzeitig* in ein Modell zur Prädiktion implementiert, was aufgrund von Kollinearitätsproblemen mit Strukturgleichungsmodellen nicht möglich ist (Hilbert et al., 2021). Um die Klausurnoten der Studierenden vorherzusagen, wurde mit einem Teil der Stichprobe ein *random forest* Modell trainiert und auf den restlichen Daten getestet. Um erfolgreiche Lernstrategien zu identifizieren, wurde das Machine Learning Verfahren *permutation feature importance* gewählt (Molnar, 2022).

Bei der Prädiktion der Klausurnoten wird einerseits vorhergesagt, ob die Klausur bestanden wurde oder nicht, und andererseits, welche Notenstufe von 1 bis 5 erreicht wurde. Trotz der für Machine Learning und Big Data verhältnismäßig kleinen Stichprobe wird bei einer 5-fachen Kreuzvalidierung eine hohe Vorhersagegenauigkeit von durchschnittlich 84% beziehungsweise 62% erreicht. Zu den

erfolgreichsten Lernstrategien gehören unter anderem das *Herausschreiben von Sätzen aus dem Vorlesungsskript* sowie die selbstberichtete Fähigkeit, *Definitionen und Sätze mit eigenen Worten erklären zu können*. Das *Auswendiglernen von Beweisen* steht dagegen in einem negativen Zusammenhang mit dem Klausurerfolg, ist also gerade *keine* adäquate Lernstrategie. Da klassische Regressionsmodelle in der Anzahl der implementierbaren Variablen beschränkt sind, ist es zwar denkbar, durch willkürliche Teilmengen, geeignete Prädiktoren zu finden, aber mühsam. Durch das Machine Learning Verfahren ist die Identifikation einfach und schnell. Darüber hinaus wird die abhängige Variable *Klausurerfolg* nach Verständnis- beziehungsweise reinen Rechenaufgaben differenziert, um aufgabenspezifische Unterschiede zu adressieren. Hierzu werden im Vortrag differentielle Effekte der Lernstrategien berichtet. Ergänzend werden Ergebnisse von Machine Learning Verfahren und klassischen Regressionen berichtet, bei denen einzelne Lernstrategien zu Skalen zusammengefasst wurden. Weiterhin wird diskutiert werden, inwieweit durch solche Verfahren nicht nur *Strategien* identifiziert werden können, die bestenfalls an Studierende weitergegeben werden können, sondern eventuell sogar *Studierende*, die weitere Unterstützung benötigen.

Die Struktur der Erfahrung von Bildung empirisch fassen: Ein phänomenologischer Zugang zu Interviews

Christine Becks[1], **Stefan Becks**[2]
[1]Universität Duisburg-Essen, Deutschland; [2]Westfälische Hochschule

Unser Beitrag stellt das phänomenologische Interview als innovative empirische Erhebungsmethode des Erlebens von institutionalisierter Bildung vor, besonders im Kontext von Schule und Hochschule. Seine Methodologie führt die Phänomenologie als philosophische Tradition ein, die es der empirischen Bildungsforschung ermöglicht, Erfahrung und deren Deutung als Grundlage des individuellen Handelns und Verstehens zu fassen und sie als Muster systematisiert zu erheben, zu verstehen und zu erklären. In dieser Weise durchdringt der phänomenologische Ansatz alle Forschungsphasen, von der Konzeptionalisierung des Forschungsvorgehens bis zur Daten-interpretation. Der Ansatz hat besonders für die Erforschung komplexer sozialer Systeme wie Schule und Hochschule vielversprechendes Potenzial: Er erfasst Erfahrung von innerhalb des Systems und gibt Aufschluss über von außen nicht sicht- oder erfahrbaren Eigenlogiken.

Um empirische Bildungsforschung multiperspektivisch und multiparadigmatisch anzulegen, ist es notwendig, eine Vielzahl an Perspektiven systematisch und methodisch gesichert einzuholen, um die komplexe soziale Wirklichkeit institutionalisierter Bildung so zu erheben und abzubilden, wie sie die Beteiligten erleben. Denn phänomenologisch gesprochen ist (Hoch)Schule das, was die Beteiligten als (Hoch)Schule erleben. Hier leistet die angewandte Phänomenologie (Zahavi 2019a, 2019b; Smith 2018) einen wesentlichen Beitrag, indem sie über die einzelne Erfahrung hinaus die Struktur der Erfahrung und ihre verschiedenen Deutungsmuster erschließt, die erklären, warum so und nicht anders – oft kontraintuitiv und entgegen eigener Vorteile – gehandelt wird.

Wir entwickeln unser Argument beispielhaft anhand des Forschungsprojekts der Westfälischen Hochschule zur Prävention von Studienabbruch, in dem die Methode des phänomenologischen Interviews (Bevan 2014) in ihrer Anwendung innerhalb institutioneller Bildung geschärft und erweitert werden soll. Seit Juli 2022 wird dort versucht, auf Grundlage von Learning Analytics (vgl. Sonderlund

et al. 2019; Lester et al. 2018) die Studienabbruchquote zu reduzieren, indem abbruchsgefährdete Studierenden identifiziert und unterstützt werden. Aktuelle Learning Analytics-Systeme können Auf Basis der vorhandenen Datenbanken (Prüfungssystem, eLearning-Plattform, etc.) einen bevorstehenden Studienabbruch vorhersagen. Dabei erlauben die verwendeten Datenquellen zwar einen Einblick in das *Was* des Studierendenhandelns, können aber kaum das *Warum* liefern. Letzteres ist aber entscheidend, um effektive Interventionen zu entwickeln. Wir gehen auf Basis des phänomenologischen Zugangs davon aus, dass alle institutionellen Lernerfahrung das sichtbare Handeln beeinflusst. Dementsprechend verstehen wir Studienabbruch als Prozess und wollen die Struktur dieser Erfahrung studierendenseitig sichtbar zu machen, sowie Ursachen für und Auslöser von Studienabbruch herausarbeiten.

Das phänomenologische Interview (Bevan 2014) besteht aus drei Teilen: Zunächst wird die Lebenswelt des/r Befragten detailreich und greifbar erfasst (*contextualizing*). Das Gespräch nähert sich dem Forschungsgegenstand dann möglichst vielschichtig, beispielhaft und vollständig (*apprehending*). Letztlich werden Klärungen, Kontrastierungen, Variationen des Erzählten nachgefragt (*clarifying*). Das ko-konstruktive Moment erscheint beim Nachfragen besonders entscheidend, da Momente des verbalisierten Nachdenkens und Schärfens hervorgebracht werden sollen. Bevan lässt die tatsächliche Gesprächsführung, die jene Offenheit und möglichen Selbstkritik systematisch anleiten soll, völlig offen. Wir greifen auf die existenzieller Gesprächsführung zurück (Kolbe 2016), die ebenfalls aus der Tradition der Phänomenologie stammt und Interviews als Gespräch Begegnung versteht, die beidseitig bedeutsam werden muss, um ebenbürtige Beteiligung und Gestaltung am Gespräch zu erwirken. So gibt es kein Objekt einer Betrachtung gibt, sondern nur Subjekte einer Begegnung. (ebd., S. 47). Gesprächsinhalte müssen existenziell bewegen, also unmittelbare Relevanz haben (ebd., S. 45f) um das ko-konstruktive Moment in der personellen Anfrage (ebd.., S. 45) gemeinsam zu entwickeln. Die Muster sind beispielsweise mit thematischer Analyse (Braun/Clarke 2006) aus den einzelnen Gesprächen und dann über die Gespräche hinweg zu entwickeln und zu sättigen.

Mit Blick auf das derzeit noch mangelnde Erklären und Verstehen von Eigenlogiken institutionalisierter Bildung erörtert unser Beitrag entlang des konkreten Forschungsprojekts eine Methode und Methodologie, die Erfahrungs- und Handlungswissen adressat*innenseitig systematisiert freilegt. Dieser innovative Zugriff ist für (Hoch)Schulforschung und deren Interventionen relevant; für das konkrete Projekt liegt die Bedeutung im Explizieren der Wirkungszusammenhänge, die verschiedene quantitative Daten durch validierte Deutungs- und Begründungsmuster erweitern.

H-06: Selbstreguliertes Lernen

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S06 S04 B06
Chair der Sitzung: Maren Stern

Behaviorale Selbstregulationsstrategien und Zielorientierungen von Kindergartenkindern beim Bearbeiten einer Problemlöseaufgabe

Maren Stern, Silke Hertel
Heidelberg University, Germany

Theoretischer Hintergrund

Selbstregulationsfähigkeiten stellen eine wichtige Voraussetzung für erfolgreiches Lernen dar (McCelland et al., 2012). Die Entwicklung zur Selbstregulation beginnt bereits in der frühen Kindheit. Hier lernen Kinder, das eigene Verhalten den kognitiven, emotionalen und sozialen Anforderungen einer Situation anzupassen (Poser & Rothbart, 2000), um zielgerichtet und kontrolliert handeln zu können (Mulder et al., 2009). Bei der Bearbeitung von Problemlöseaufgaben spielen kognitive, metakognitive und motivationale Strategien eine zentrale Rolle (Dermitzaki et al., 2009). Bislang wurden diese Strategien insbesondere vom Grundschulalter an untersucht (Roebers et al., 2014) und es existieren nur wenige Studien zum Kindergartenalter (Zhang & Whitebread, 2017).

Neben der Frage, welche konkreten Strategien Individuen zur Zielerreichung nutzen, ist die motivationale Ausrichtung auf konkrete Ziele ein weiterer wichtiger Bestandteil für erfolgreiches Lernen (Dweck & Legget, 1988). In der Literatur wird Zielorientierung in Lern- und Leistungsziele unterteilt. Dabei gehen Lernziele mit positiven lern- und leistungsbezogenen Emotionen und Kognitionen sowie gute Leistungen einher, während für Leistungsziele eher eine uneindeutige Befundlage vorliegt (Elliot & Hulleman, 2017). Wie Zielorientierungen und Selbstregulationsstrategien bei Kindergartenkindern zusammenspielen wurde in der Forschung bisher nicht adressiert.

Fragestellungen

(1) Wie hängen behaviorale Selbstregulationsstrategien und Problemlöseleistung von Kindergartenkindern zusammen?

(2) Ist die Zielorientierung von Kindern mit deren behavioralen Selbstregulationsstrategien und Problemlöseleistung assoziiert?

Methode

Daten von N=144 Kindern im Alter von vier und fünf Jahren (M=4.55, SD=0.54) wurden im Labor erhoben. Mittels Tablet-PC wurde die Zielorientierung (Lern- vs. Leistungsziel, in Anlehnung an Gunderson et al., 2013) erfasst. Anschließend bearbeiteten die Kinder eigenständig eine Problemlöseaufgabe (Melchers & Melchers, 2015). Die Kinder wurden videografiert und deren Problemlöseleistung sowie kognitive, metakognitive und motivationale Selbstregulationsstrategien anhand eines adaptierten Kodierschemas erfasst (Dermitzaki et al., 2009).

Ergebnisse

Zur Beantwortung der ersten Fragestellung wurde eine hierarchische Regressionsanalyse zur Vorhersage der Problemlöseleistung gerechnet. Im ersten Schritt erklärten Alter ($ß$=.54, p<.001) und Geschlecht ($ß$=.06, p=.404) des Kindes 29% der Varianz in der kindlichen Problemlöseleistung (F(2,144)=29.52, p<.001). Die Hinzunahme der kognitiven ($ß$=.27, p<.001) und metakognitiven

Strategien (ß=.31, *p*<.001) im zweiten Schritt ging mit einer substantiellen Verbesserung in der Varianzaufklärung einher (ΔR^2=.21, *F*(4,144)=35.85, *p*<.001). Das Alter des Kindes blieb signifikanter Prädiktor (ß=.30, *p*<.001). Die Hinzunahme der motivationalen Strategien (ß=.26, *p*<.001) führte zu einer weiteren Verbesserung in der Varianzaufklärung (ΔR^2=.05, *F*(5,144)=35.13, *p*<.001). Alter (ß=.22, *p*<.001), kognitive (ß=.23, *p*=.004) und metakognitive Strategien (ß=.27, *p*<.001) des Kindes blieben signifikante Prädiktoren.

Zur Beantwortung der zweiten Fragestellung wurde zunächst überprüft, ob sich die Kinder in Abhängigkeit ihrer Zielorientierung in der Anwendung der Selbstregulationsstrategien und in der Problemlöseleistung unterscheiden. Die Analysen ergaben, dass Kinder mit einer Lernzielorientierung mehr kognitive (*t*(143)=2.20, *p*=.030, *d*=0.64), metakognitive (*t*(143)=2.38, *p*=.019, *d*=0.65) und motivationale Strategien (*t*(143)=2.53, *p*=.013, *d*=0.65) nutzen sowie eine bessere Problemlöseleistung (*t*(143)=3.72, *p*<.001, *d*=3.40) erzielen als Kinder mit einer Leistungszielorientierung. Anschließend wurde mittels Pfadanalysen untersucht, ob der Effekt der Zielorientierung über die Selbstregulationsstrategien vermittelt wird. Es zeigte sich, dass die motivationalen Strategien den Zusammenhang zwischen Zielorientierung und Problemlöseleistung mediieren (*ab*=.04, BCa CI (0.008, 0.557)), während kein signifikanter Mediatoreffekt für die kognitiven (*ab*=.03, BCa CI (-0.044, 0.545)) und metakognitiven Strategien (*ab*=.04, BCa CI (-0.022, 0.610)) des Kindes gefunden wurde.

Die Befunde weisen darauf hin, dass sowohl kognitive, metakognitive als auch motivationale Strategien zur Vorhersage der kindlichen Problemlöseleistung beitragen. Während bisherige Studien metakognitive Strategien als Hauptprädiktor für kindliches Problemlösen identifizieren (z.B. Zhang & Whitebread, 2017), betont die vorliegende Studie neben den kognitiven auch die motivationalen Strategien. Dabei scheinen motivationale Strategien den positiven Effekt der Lernziele auf die Problemlöseleistung des Kindes zu vermitteln.

Als Limitation der Studie ist zu beachten, dass die Zielorientierung des Kindes dichotom erhoben wurde und weitere Differenzierungen (z.B. Annäherung vs. Vermeidung) vernachlässigt wurden. Insgesamt haben die Ergebnisse wichtige Implikationen für die Förderung von Selbstregulation im frühen Kindesalter und zukünftige Forschung.

Aspekte der Nutzung eines Webtools zur Förderung selbstregulierten Lernens an der Hochschule – Welche Studierende nutzen ein solches Webtool?

Maximilian Pfost, Peter Kuntner, Simone A. Goppert, Vanessa Hübner
Otto-Friedrich-Universität Bamberg, Germany

Theoretischer Hintergrund

Aufgrund seiner vielen Freiheiten mit Bezug auf die Ausgestaltung des Studiums und der Möglichkeiten der Wahrnehmung seiner Angebote stellt die Hochschule im Vergleich zur Schule deutlich höhere Anforderungen an die Regulation des individuellen Lernverhaltens (Goppert, Neuenhaus & Pfost, 2021). Um die Studierenden in der Regulation ihres Lernens zu unterstützen haben wir ein Webtool entwickelt und vorlesungsbegleitend eingesetzt. In diesem Webtool sind Studierende im wöchentlichen Rhythmus angehalten, ihr Lernverhalten und ihre Kenntnisse einzuschätzen sowie Fragen zum Inhalt der Lehrveranstaltung zu beantworten. Das Webtool bietet

den Studierenden darüber hinaus die Möglichkeit, Feedback, beispielsweise zur Anzahl (richtig) bearbeiteter Aufgaben, abzurufen. Daneben wurden Gamification-Elemente zur Motivationsaufrechterhaltung integriert.

Die Nutzung des Webtools erfolgt auf freiwilliger Basis, implizit beruhend auf der Annahme, dass sich die Studierenden ihrer Lernbedürfnisse ausreichend bewusst sind, beziehungsweise über die hierfür notwendigen metakognitiven Fähigkeiten verfügen (vgl. Clarebout, Elen, Collazo, Lust & Jiang, 2013). Vorangehende Studien zeigen jedoch, dass Lernende elektronische Unterstützungsangebote vielfach nicht angemessen oder nicht in dem Umfang nutzen, wie es von den Lehrenden intendiert wird (Jansen, van Leeuwen, Janssen, Conijn & Kester, 2020; Lust, Elen & Clarebout, 2013). Da implizit anzunehmen ist, dass elektronische Unterstützungsangebote für das Lernen nur ihre Wirksamkeit entfalten können, wenn diese auch – zumindest in gewissem Umfang – genutzt werden, erhält die Forschungsfrage nach Variablen, die mit dem Umfang und der Bereitschaft zur Nutzung eines Webtools in Zusammenhang stehen, hohe Relevanz.

Fragestellung:

Welche Merkmale der Studierenden, insbesondere mit Bezug auf das selbstregulierte Lernen, stehen in Zusammenhang mit der Nutzung eines, die Lehrveranstaltung ergänzenden, Webtools?

Methode:

Die Ergebnisse beruhen sowohl auf Fragebogendaten als auch auf Daten zur Nutzung des Webtools von insgesamt 223 Studierenden. Neben Variablen zur Webtool-Nutzung wurde unter anderem das Geschlecht, die Abiturnote der Teilnehmenden sowie Einschätzungen zum selbstregulierten Lernen (Planung; Überwachung; Regulation; Zeitmanagement) erfasst und miteinander in Beziehung gesetzt. Die Fragebogendaten beziehen sich auf die Gruppe der Nutzerinnen und Nutzern des Webtools, aber auch auf Studierende, die das Webtool nicht genutzt haben.

Ergebnisse:

Die Auswertungen zeigen, dass weniger als die Hälfte der Studierenden das Webtool zu nutzen scheinen. Statistisch signifikante Zusammenhänge konnten dabei mit den Variablen *Geschlecht, Abiturnote* sowie mit der *Intention zur Planung der Nachbereitung der Lehrveranstaltung* gefunden werden. So nutzten mehr weibliche als männliche Studierende das Webtool. Nutzerinnen und Nutzer hatten darüber hinaus im Durchschnitt eine bessere Abiturnote und planten auch im Vorfeld häufiger, die Lehrveranstaltung nachzubereiten. Interessant war jedoch auch, dass beispielsweise das selbst eingeschätzte Zeitmanagement der Studierenden nicht mit deren Webtool-Nutzung im Zusammenhang stand. Die Befunde zeigen auch, dass die Mehrheit der Nutzerinnen und Nutzer das Webtool als lernförderlich betrachten.

Diskussion:

Die Ergebnisse zeigen, dass durch ein freiwillig ergänzendes Webtool nur ein Teil der Studierenden erreicht wird, wobei weibliche Studierende und Studierende mit besserer Abiturnote leichter erreichbar scheinen. Gerade der Zusammenhang mit der Abiturnote ist hierbei interessant, da Schulnoten nicht nur ein guter Prädiktor für den Studienerfolg sind (Trapmann, Hell, Weigand &

Schuler, 2007), sondern auch mit dem Ausmaß an Selbstregulation und Selbstkontrolle im Studium in Zusammenhang stehen (Galla et al., 2019). Der gefundene Zusammenhang mit der Intention der Studierenden zur Nachbereitung der Lehrveranstaltung deutet auf die Relevanz von Planungsprozessen für die Nutzung des Webtools hin. Die Ergebnisse werden vor dem Hintergrund der Verbesserung des Lehrangebots mittels Webtools in ihren Implikationen diskutiert.

Is your mind set? Mindsets von Schüler:innen und Lehrpersonen sowie deren Zusammenspiel

Silke Hertel[1], Yves Karlen[2]
[1]Universität Heidelberg, Germany; [2]Pädagogische Hochschule FHNW, Switzerland

Mindsets und Überzeugungen von Schüler:innen und Lehrpersonen nehmen im Schulalltag eine wichtige Rolle ein. Unser Beitrag knüpft an aktuelle Forschung zu Mindsets und selbstreguliertem Lernen (SRL) an, die die Bedeutung von Growth Mindsets zu Intelligenz (INTGM) und zu SRL (SRLGM) für das SRL und akademische Leistungen betont (Burnette et al., 2013; Chen & Tutwiler, 2017; Dweck & Master, 2008; Hertel & Karlen, 2020; Sisk et al., 2018; Yu & McLellan, 2020) und einen Zusammenhang zu Überzeugungen und Unterrichtspraktiken der Lehrpersonen herstellt (Bostwick et al., 2020; Karlen et al., 2020; Martin et al., 2021; Yu et al., 2022). Bislang gibt es allerdings nur sehr wenige schulklassenbasierte Studien, die den Zusammenhang der Growth Mindsets von Schüler:innen und Lehrpersonen sowie der SRL Förderung im Unterricht betrachten.

In dem Beitrag werden die folgenden drei Fragestellungen untersucht:

FF1: Stehen SRLGM und INTGM von Schüler:innen in einem positiven Zusammenhang mit ihren schulischen Leistungen?

FF2: Stehen SRLGM und INTGM der Schüler:innen in einem positiven Zusammenhang mit der wahrgenommenen SRL-Förderung im Unterricht sowie mit den SRLGM und INTGM der Lehrpersonen?

FF3: Wird der Zusammenhang von SRLGM und INTGM und Leistung über das schulische Selbstkonzept und das metakognitive Strategiewissen der Schüler:innen vermittelt?

Die Analysen basieren auf Daten von *N=1468* Schüler:innen aus *N=117* Klassen sowie ihren Lehrpersonen. Mittels Fragebögen wurden SRLGM, INTGM, schulisches Selbstkonzept, wahrgenommene SRL-Förderung im Unterricht sowie als Kontrollvariablen Geschlecht, Alter, Migrationshintergrund und sozioökonomischer Status (SES) der Schüler:innen erfasst. Zudem wurde ein Test zur Erfassung des metakognitiven Strategiewissens eingesetzt und die schulischen Leistungen (Zeugnisnote) erhoben. Auf Ebene der Lehrpersonen wurden SRLGM, INTGM, Erfahrung in der SRL-Förderung im Unterricht, Selbstwirksamkeit in der Förderung und Diagnostik von SRL, sowie Geschlecht und Diensterfahrung mittels Fragebogen erfasst. Die Analyse der Daten erfolgte im Rahmen von zwei multi-level Regressionsmodellen (MSEM) mit MPlus. Das erste Modell zielte auf die Beantwortung von FF1 und FF2. Im zweiten Modell wurden zusätzliche Variablen aufgenommen, um den Mediationseffekt aus FF3 zu untersuchen.

In Modell 1 (Chi^2= 61.668, df=17, RMSEA=.042, CFI=.88, $SRMR_{within}$= .038, $SRMR_{between}$= .082) zeigt sich ein positiver Zusammenhang zwischen dem SRLGM der Schüler:innen und ihrer schulischen Leistungen (FF1). Zudem zeigen sich signifikante Zusammenhänge zum Geschlecht der

Schüler:innen und der von den Schüler:innen wahrgenommenen SRL-Förderung. Bei der Betrachtung auf Klassenebene zeigte sich ein positiver Zusammenhang zwischen der Erfahrung der Lehrperson in der SRL-Förderung und dem SRLGM der Lehrpersonen (FF2). Zudem zeigt sich für das INTGM der Schüler:innen der Klasse ein signifikant positiver Zusammenhang zum INTGM der Lehrpersonen. Die wahrgenommene SRL-Förderung der Schüler:innen in der Klasse zeigt einen positiven Zusammenhang zur Erfahrung der Lehrperson in der SRL-Förderung und negative Zusammenhänge zum mittleren Alter der Schüler:innen sowie zum mittleren SES. In Modell 2 (Chi^2= 107.737, df=30, RMSEA=.042, CFI=.91, $SRMR_{within}$= .033, $SRMR_{between}$= .097) wurden auf Schüler:innenebene das schulische Selbstkonzept der Schüler:innen sowie ihr metakognitives Wissen mit aufgenommen. Dabei zeigt sich, dass der Effekt des SRLGM auf die schulischen Leistungen vollständig über das schulische Selbstkonzept und das metakognitive Wissen mediiert wird (FF3). Weiterhin zeigen sich positive Zusammenhänge des schulischen Selbstkonzepts mit dem SRLGM und INTGM der Schüler:innen sowie zwischen dem metakognitiven Strategiewissen der Schüler:innen, ihrem SRLGM und INTGM und der wahrgenommenen SRL-Förderung. Auf Klassenebene zeigt sich für die Lehrpersonen ein signifikanter Zusammenhang der Erfahrung in der SRL-Förderung und der Selbstwirksamkeit in der Förderung und Diagnostik von SRL.

Die Ergebnisse unterstreichen die Bedeutung von SRLGM im Zusammenspiel mit schulischem Selbstkonzept und metakognitivem Wissen der Schüler:innen für die schulischen Leistungen. Darüber hinaus zeigen sie, dass die SRL-Förderung im Unterricht einen Ansatzpunkt zur Unterstützung günstiger Überzeugungsmuster darstellt. Auch wenn der Beitrag nur auf querschnittlichen Daten basiert und keine kausalen Aussagen möglich sind, liefert er wichtige Impulse für die weitere Forschung und die Schulpraxis.

Wer hält durch? Vorhersage der Persistenz in non-formalen Online-Kursen

Philipp Handschuh[1], Maria Klose[1], Diana Steger[2]
[1]LIfBi – Leibniz-Institut für Bildungsverläufe, Germany; [2]Universität Kassel, Germany

Non-formale Weiterbildungsmöglichkeiten bilden einen integralen Bestandteil der individuellen freiwilligen Erwachsenenbildung. Jedoch sind gerade bei diesen Kursangeboten oft die Abschlussquoten der Teilnehmenden eher gering (bspw. für MOOCS ist diese Quote meistens unter 10%; Jordan, 2014). Aufgrund der geringeren Unterstützung und Anleitung in Online-Kursen, ist Selbstregulationsfähigkeit entscheidend um die persönlichen Lernziele zu erreichen (Kizilcec et al., 2016). Dabei spielen lernpsychologisch relevante Faktoren sowie Eigenschaften der Lernumwelt eine entscheidende Rolle für das Bearbeitungsverhalten der Lernenden – v.a. die Fähigkeit des selbstregulierten Lernens (Pintrich, 2000). Der Prozess der Selbstregulation wird durch mehrere Faktoren beeinflusst, wie die Einschätzung der Werte der Aufgabe, des Nutzen des Lernens und das persönliche Interesse an dem Thema (Schunk, 2015). Dies zeigt sich auch in dem Zusammenhang der Selbstregulationsfähigkeit mit Abschließen von MOOCs (Reparaz et al., 2020). Darüber hinaus ist die intrinsische Motivation ein Prädiktor für das Bearbeitungsverhalten von Kursen, da Lernende die es schaffen ihre Motivation aufrechtzuerhalten auch in ihrem Lernverhalten persistieren (Deci, 1971). So sagt die Motivation nicht nur das Bearbeitungsverhalten (Badali et al., 2022), sondern auch die Einschreibewahrscheinlichkeit vorher (Kalenda & Kočvarová, 2022). Darüber hinaus sind die

Eigenschaften der Lernumgebung non-formalen Lernens relevant. So empfehlen Jones et al. (2021), dass Online-Kurse das Interesse der Studierende an den Kursaktivitäten wecken sollten, angemessen schwierig gestaltet werden und einen adäquaten Bearbeitungsumfang haben sollten (vgl. Jordan, 2015).

Ziel der Studie ist es das Bearbeitungsverhalten von Lernenden durch lernpsychologische Personeneigenschaften und Kurscharakteristiken vorherzusagen. Als abhängige Variable wird hierbei einerseits das Bearbeitungsverhalten betrachtet, d.h. die Anzahl der korrekt gelösten Aufgaben innerhalb eines Kurses, und andererseits über den Erhalt einer Teilnahmebestätigung. Letztere wird durch das Erreichen einer kursspezifischen Grenze von gelösten Aufgaben erlangt. Es wird erwartet, dass die Selbstregulationsfähigkeit, sowie die situative intrinsische Motivation positiv mit dem Bearbeitungsverhalten zusammenhängen. Aufgrund der Möglichkeit eigenes Wissen zu testen und somit Ziele abzugleichen (vgl. Butler & Winne, 1995), sollte auch die Anzahl der Tests innerhalb eines Kurses positiv mit dem Verhalten zusammenhängen. Lediglich die Kurslänge sollte zu einer geringeren Kursbearbeitung führen. Darüber hinaus wird das Bearbeitungsverhalten mit der individuellen Einschreibeintention der Lernenden in Zusammenhang gebracht und näher beschrieben.

Die Online-Befragung fand von Oktober 2021 bis Juni 2022 im Rahmen des digitalen Kursangebotes der OPEN vhb (Virtuelle Hochschule Bayern, 2022) statt und umfasst 51 Online-Kurse, die ohne vorgegebene Lernzeiten, kostenfrei nutzbar und frei zugänglich sind. Insgesamt haben 1.414 Lernende an der Befragung teilgenommen (Alter: M = 36.54, SD = 14.08, Range [16; 84]; davon 68% weiblich). Dabei wurden die relevanten Eigenschaften der Lernenden mit Kurzskalen sowie Subskalen etablierter Erhebungsinstrumente zu Beginn der Kursbearbeitung erhoben (situative Motivation: SIMS; Guay et al., 2000 ; Selbstregulation: über die Subskala des LIST-K „Strategien des Managements interner Ressourcen“; Klingsieck, 2018). Zur Beantwortung der Frage, wie sich die personenbezogenen und kursbezogenen Charakteristiken auf das Bearbeitungsverhalten auswirken, werden multiple lineare, sowie logistische Regressionen berechnet.

Vorläufige Ergebnisse zeigen erwartungsgemäße Zusammenhänge von Selbstregulation (positiv) und Kursdauer (negativ) mit dem Erreichen einer Teilnahmebestätigung. Der Einfluss der Testanzahl scheint vernachlässigbar zu sein. Überraschenderweise zeigt die intrinsische Motivation einen tendenziell negativ Zusammenhang mit dem Erreichen der Teilnahmebestätigung auf, wohingegen die extrinsische Motivation einen größeren positiven Effekt hat. Auf Grundlage der Ergebnisse sowie der Berücksichtigung unterschiedlicher Definitionen von Bearbeitungsfortschritt (Teilnahmebestätigung vs. gelöste Tests), sollen Ansatzpunkte für individuelle Lerndesigns gefunden werden, um die Bearbeitungsbereitschaft weiter zu erhöhen. In weiteren Analysen sollen das Bearbeitungsverhalten und die Erreichung der Teilnahmebestätigung verglichen werden, um letztendlich den „erfolgreichen Lernenden“ in selbstgesteuerten Online-Kursen zu beschreiben und zu diskutieren. Diese Diskussion wird um die Betrachtung der Einschreibeintention ergänzt, bei der sich zwischen Personen mit und ohne Teilnahmebestätigung Unterschiede in den Zielen „Um eine Teilnahmebestätigung zu erhalten“ und „Zum Spaß oder als Herausforderung“ (häufiger bei Personen ohne Teilnahmebestätigung) andeuten.

H-07: Abitur und automatisches Scoring

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S05 T00 B59
Chair der Sitzung: Alexander Groß

„…dass es nicht heißt, hier kriegt man das Abitur hinterhergeworfen und woanders ist es so schwer.“ Rationalitätsmythen als Grundlage bildungspolitischer Reformentscheidungen?

Alexander Groß
Universität Koblenz, Deutschland

Theoretischer Hintergrund

Bildungspolitische Entscheidungen der Bundesländer und insbesondere die Implementation von Neuerungen stehen immer im Kontext einer nachzuweisenden Verbesserung des Status quo und werden entsprechend bewertet. Wird man als Land jenem hier entstehenden gesellschaftlichen Rechenschaftsdruck und den in diesem Kontext formulierten Erwartungen nicht gerecht, kann der gesamten Art und Weise der Schulorganisation Legitimität (z.B. Koch, 2018) abgesprochen werden, woraus sich aus der Perspektive des soziologischen Neo-Institutionalismus (z.B. Walgenbach & Meyer, 2008) der legitimatorisch-defizitäre Länderstatus „debated“ (Deephouse et al., 2017, S. 33), also in Frage gestellt, ergibt. Konkret wird durch ein solch attestiertes ‚performance‘ Defizit (vgl. Oliver, 1992) die eigene Reputation individuell und im Bundesländervergleich potentiell negativ beeinflusst. Ob die hier zentralen „Annahmen über erstrebenswerte Zielsetzungen und zweckmäßige Mittel der Zielerreichung“ (Herbrechter, 2018, S. 26) eine tatsächlich rationale Grundlage haben, spielt indes keine Rolle; Rationalität generiert sich nicht durch deren Objektivität und Überprüfbarkeit, sondern dadurch, dass es in der gesellschaftlichen Umwelt einen *Glauben* an eine entsprechende Wirksamkeit und Angemessenheit gibt; so geht es also nicht um gesichertes Wissen, sondern lediglich um die „supposition“ (Meyer & Rowan, 1977, S. 34) und eine daraus resultierende (De-)Legitimation einer Neuerung. Es ist wahrscheinlich, dass solche Rationalitätsmythen im Sinne ‚vermeintlichen Wissens‘ regelmäßig auch bildungspolitische Entscheidungen (mit)beeinflussen; gleichwohl ist diese Form der legitimatorischen Handlungsgrundlage bis dato nicht systematisch erforscht. Dieser Beitrag stellt Ergebnisse eines Forschungsprojekts zu diesem Thema am Reformbeispiel der Einführung ländergemeinsamer Abiturprüfungsaufgaben vor.

Fragestellungen

1. Welche Rationalitätsmythen werden bei der (De-)Legitimation ländergemeinsamer Abiturprüfungsaufgaben auf Landesebene sichtbar?

2. Inwiefern stellen diese eine tatsächliche bildungspolitische Entscheidungsgrundlage dar?

Methode

Die hier berichteten Ergebnisse sind ein Ausschnitt aus einer qualitativ-explorativen Interviewstudie zum gesamten Implementationsprozess ländergemeinsamer Abiturprüfungsaufgaben im schulischen Mehrebenensystem (vgl. Autor, 2022). Bei der Stichprobenziehung wurden im Sinne eines „intensity samplings“ (Patton, 1990, S. 171) insgesamt drei Bundesländer, mit dem Fokus auf möglichst heterogene Fälle, in die Analyse einbezogen. Ziel war es, jeweils die entsprechenden Gelingensbedingungen für einen bildungspolitischen Reformerfolg in diesem Kontext offenzulegen. Dieser Beitrag stellt die Ergebnisse in Bezug auf konkret nachweisbare Rationalitätsmythen als Handlungsgrundlage im Umgang mit dieser Neuerung vor und gegenüber. Dabei wurden jeweils zwei

ministerial für die Abiturprüfung verantwortliche Akteure sowie drei Gymnasien pro Land und dort wiederum jeweils zwei Lehrkräfte mit langjähriger Abiturprüfungserfahrung pro Schule als Expert*innen interviewt (z.B. Bogner et al., 2014) und in die Analyse einbezogen (n=20). Die Auswertung des Materials erfolgte in Form einer inhaltlich strukturierenden qualitativen Inhaltsanalyse (Kuckartz, 2016); die Kategorien zur Kodierung der Interviewdaten wurden in einem (theoriebasiert) deduktiv- (datengesteuert) induktiven Verfahren entwickelt (vgl. Döring & Bortz, 2016, S. 542), die Kodierungen selbst konsensuell validiert (vgl. Bortz & Döring, 2006, S. 328).

Ergebnisse

Es zeigt sich, dass, trotz unterschiedlicher Prüfungstraditionen, Rationalitätsmythen in allen betrachteten Ländern eine wesentliche legitimatorische Rolle spielen und ebenso eine zentrale bildungspolitische Handlungsgrundlage darstellen. Dabei ist solch vermeintliches Wissen insbesondere in Bezug auf die Einschätzungen zur Qualität des eigenen Abiturs und der gestellten Prüfungsanforderungen nachweisbar: In *Land 1* verspricht man sich von der Einführung gemeinsamer Prüfungsaufgaben eine Ablösung des extern entstandenen Mythos eines minderwertigen eigenen Abiturs, entsprechend reformfreundlich ist die vorherrschende Einstellung; in *Land 2* reproduziert man den Mythos einer qualitativ überlegenen dezentral organisierten Abiturprüfung und de-legitimiert so die ländergemeinsamen Aufgaben bereits als Instrument; in *Land 3* bewertet man das eigene Abitur als bis dato bundesweit am anspruchsvollsten und möchte diesen auch extern zugewiesenen mythischen Status qualitativer Überlegenheit weiterhin erhalten.

Durch die Herausarbeitung potentiell nachweisbarer Rationalitätsmythen und der damit einhergehenden Erweiterung des Analysefokus über die ausschließlich inhaltliche Ebene einer Neuerung hinaus, wird allgemein der häufig virulenten Multikausalität einer Ablehnungs- bzw. Akzeptanzentscheidung auch analytisch stärker Rechnung getragen. Dies trägt des Weiteren dann auch dazu bei, die Chancen für einen tatsächlichen Reformerfolg letztlich fundierter einschätzen zu können.

Die Abiturprüfungen ausfallen lassen? Zur Passung von Kurs-und Prüfungsnoten auf Ebene der Gesamtqualifikation

Marko Neumann[1], Michael Becker[2,1], Jürgen Baumert[3], Olaf Köller[4], Kai Maaz[1]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main; [2]Technische Universität Dortmund; [3]Max-Planck-Institut für Bildungsforschung, Berlin; [4]IPN-Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel

Die (mangelnde) Vergleichbarkeit von Noten und Abschlusszertifikaten hinsichtlich dahinterstehender Kompetenzen und Fähigkeiten ist ein Dauerthema in der Diskussion um Verteilungsgerechtigkeit. Dies ist wenig verwunderlich, sind an erreichte Zertifikate und Noten unmittelbar Ausbildungs-, Berufs- und damit auch Lebenschancen geknüpft.

Ein besonderer Kristallisationspunkt der Debatten um mangelnde Vergleichbarkeit und unterschiedliche Bewertungsmaßstäbe in der Zertifikats- und Notenvergabe ist das Abitur. Hier stehen zum einen immer wieder Unterschiede in den Leistungsanforderungen und Bewertungsmaßstäben zwischen den Bundesländern im Fokus. Zum anderen wird ein im Zeitverlauf abnehmendes Anforderungsniveau beim Abitur konstatiert, welches sich in gestiegenen

Abiturgesamtnoten zeige, ohne dass dieser Anstieg durch entsprechende Kompetenzstände der im Zuge der Bildungsexpansion stark angestiegenen Zahl von Abiturient*innen hinterlegt sei (Stichwort „Noteninflation“) (Brodkorb & Koch 2020).

Die Frage der Leistungs- und Prüfungsanforderungen bei Abitur kam auch in der Corona-Pandemie eine hohe Relevanz zu. In Zeiten von Homeschooling und schwer kalkulierbaren Infektionsrisiken wurde dabei von einigen Akteuren die Durchführung der Abiturprüfungen als Ganzes zur Disposition gestellt. Stattdessen wurde für ein so genanntes „Anerkennungsabitur“ geworben, welches ausschließlich die in den vier Schulhalbjahren der Qualifikationsphase der Oberstufe erzielten Kursergebnisse für die Gesamtqualifikation (Abiturgesamtnote) zu Grunde legt (Hoffmann et al., 2022a). Ein Abitur ohne Abiturprüfungen – ein in der deutschen Abiturtradition unter normalen Umständen nur schwer vorstellbarer Gedanke: Die Bewährung in einer abschließenden Prüfungssituation ist seit jeher fester Bestandteil des Abiturs und die im Prüfungsblock des Abiturzeugnisses erreichten Noten tragen ein Drittel zur Abiturgesamtnote bei. Die Abiturprüfungen stellen ein wichtiges Instrument in den Bemühungen um einheitliche Leistungsanforderungen und vergleichbare Bewertungsmaßstäbe dar (Hoffmann et al., 2022b; Klieme, 2022).

Die KMK hat sich letztlich für die Durchführung der Abiturprüfungen entschieden, nicht zuletzt um dem potentiellen Makel eines Anerkennungsabiturs mit geringeren Anforderungen aufgrund fehlender Prüfungen vorzubeugen. Kritiker eines Anerkennungsabiturs äußerten in diesem Zusammenhang auch die Befürchtung eines starken Anstiegs der Abiturdurchschnittsnoten und eine mangelnde Vergleichbarkeit der Reifezeugnisse zu den vorangegangenen Jahrgängen. Die Befürchtung eines Notenanstiegs bei alleiniger Berücksichtigung der Kursnoten impliziert einerseits die Möglichkeit eines höheren Schwierigkeitsniveaus der Abiturprüfungen im Vergleich zu den Kursergebnissen und/oder andererseits strengere Bewertungsmaßstäbe bei den Abiturprüfungen. Die bislang wenigen vorliegenden Forschungsarbeiten (vgl. im Überblick Hoffmann et al., 2022a) deuten in der Tat auf tendenziell günstigere Kursnoten im Vergleich zu den Prüfungsnoten hin. Die Befundlage ist jedoch keinesfalls einheitlich und insgesamt als sehr überschaubar einzustufen.

An dieser Stelle setzt der vorliegende Beitrag an, in dessen Zentrum die Frage nach der Passung von Kurs- und Prüfungsnoten auf der globalen Ebene der Gesamtqualifikation im Abiturzeugnis steht. Konkret soll untersucht werden, inwieweit der Prüfungsnotenblock in der Gesamtqualifikation (maximal 300 Punkte) im Mittel höher, niedriger oder gleich hoch ausfällt wie der Kursnotenblock (maximal 600 Punkte). Zum Zwecke der Vergleichbarkeit werden die Ergebnisse aus beiden Blöcken auf eine gemeinsame Metrik gebracht. Neben dem mittleren Vergleich soll auch die Streuung der Differenz von Kurs- und Prüfungsnoten untersucht werden, um Auskünfte darüber zu erhalten, welche Schüler*innenanteile von einer Beschränkung auf die Kursergebnisse profitierten und für welche Anteile sich Nachteile ergäben. Der Beitrag führt erste vorliegende Analysen (vgl. Hoffmann et al., 2022) fort, in denen auf der Datengrundlage der BERLIN-Studie zur Evaluation der Berliner Schulstrukturreform (Neumann et al., 2017) für den Abiturjahrgang 2014 an allgemeinbildenden Gymnasien ein Vergleich von Kurs- und Prüfungsnotenblock vorgenommen wurde. Die Analysen ergaben im Mittel niedrigere Prüfungsergebnisse im Vergleich zu den Kursergebnissen. Rund zwei Drittel der Schüler*innen würden von einer Beschränkung auf die Kursergebnisse profitieren, für ein

Drittel würde die Beschränkung eine schlechtere Abiturgesamtnote nach sich ziehen. Im vorliegenden Beitrag sollen die Analysen auf den Berliner Abiturjahrgang 2017 sowie alternative Wege zur Hochschulreife (insbesondere Integrierte Sekundarschule und Berufliche Gymnasien) ausgeweitet werden, um die Befundlage zur Passung von Kurs- und Prüfungsergebnissen auf eine breitere empirische Basis zu stellen.

„Was ist das Abitur noch wert?“: Eine Diskursanalyse der öffentlichen Debatte um Qualität und Vergleichbarkeit der Allgemeinen Hochschulreife

Annemarie Müller, Svenja Mareike Schmid-Kühn
Universität Koblenz, Deutschland

Die Allgemeine Hochschulreife (AHR) stellt als höchster Schulabschluss in Deutschland die formale Berechtigung „zum Studium aller Studiengänge an allen Hochschulen der Bundesrepublik“ (Neumann & Trautwein 2019, S. 533) dar. Hinsichtlich ihrer Qualität und Vergleichbarkeit ist die AHR „Gegenstand wiederkehrender Diskussionen in der (Bildungs-)Politik und Administration, in der Schulpraxis, der Wissenschaft und nicht zuletzt auch der breiten Öffentlichkeit“ (Hoffmann et al. 2022, S. 7). Insbesondere zeigt sich eine „beständige, bis heute andauernde Debatte um die vermeintlich mangelnde Vergleichbarkeit der Abiturprüfungen der Länder sowie der dabei erzielten Noten“ (Hoffmann, Schröter & Stanat 2022, S. 40), verbunden mit der Frage nach Gerechtigkeit bei der Vergabe von Studienplätzen. Dies hat zu einer Reihe von bildungspolitischen Maßnahmen geführt, mit dem Ziel, die Vergleichbarkeit der AHR zu erhöhen. Hierzu zählen vor allem die Einführung zentraler Abiturprüfungen in allen deutschen Bundesländern mit Ausnahme von Rheinland-Pfalz (vgl. ebd., S. 44), die Erarbeitung von Bildungsstandards (vgl. ebd., S. 47) sowie die schrittweise Implementation ländergemeinsamer Aufgabenpools für ausgewählte Fächer (vgl. ebd., S. 49ff.). Trotz der bisherigen Maßnahmen sind Vergleichbarkeit und Qualität der AHR nach wie vor Gegenstand öffentlicher bzw. medialer Diskussionen, wobei jedoch häufig offen bleibt, welches Verständnis z.B. von Vergleichbarkeit im Zusammenhang mit der AHR überhaupt zugrunde gelegt wird (vgl. Autor*in 2022a). An dieser Stelle setzt das im Rahmen des Beitrags vorgestellte Forschungsprojekt an, welches aus einer diskursanalytischen Perspektive folgende Fragestellung zu beantworten versucht: Welche diskursiven Strukturen kennzeichnen die öffentliche Debatte um Vergleichbarkeit und Qualität der Allgemeinen Hochschulreife? Zur Untersuchung dieser Fragestellung wird das Forschungsprogramm der Wissenssoziologischen Diskursanalyse (WDA) nach Reiner Keller (2011) herangezogen. Die WDA zielt auf die Rekonstruktion von „Prozessen und Praktiken der Produktion und Zirkulation von Wissen“ (Keller 2011, S. 192f.), untersucht also, unter welchen Umständen gesellschaftliche Wissensbestände hergestellt und verbreitet werden. Sie versteht Diskurse als „durch Regeln strukturierte serielle Praktiken des Zeichengebrauchs in sozialen Arenen“ (Keller 2019, S. 44). Zur Untersuchung öffentlicher Diskurse wie der Debatte um Qualität und Vergleichbarkeit der AHR eignen sich insbesondere die Massenmedien, da sie als „zentrale Arenen der gesellschaftlichen Wirklichkeitskonstruktion“ (Keller 2015, S. 29) gelten und somit gesellschaftliche Wissensbestände entscheidend prägen. Den Untersuchungskorpus bilden 3989 Zeitungsartikel überregionaler Tages-, Sonntags- und Wochenzeitungen aus dem Zeitraum von 1970 bis 2019, ausgehend von der Neugestaltung der gymnasialen Oberstufe im Jahr 1972 als einem zentralen Ereignis im Hinblick auf

die Entwicklung der AHR (vgl. Klemm 2022, S. 34). In einem ersten Analyseschritt wurden die Artikel zahlenmäßig ihrem Veröffentlichungsjahr zugeordnet, um einen Überblick über den Gesamtverlauf des Diskurses innerhalb der letzten fünfzig Jahre zu erhalten (vgl. Autor*in 2022b). Weiterhin wurden fünf zentrale Themenstränge identifiziert sowie deren Präsenz im zeitlichen Verlauf des Diskurses untersucht (vgl. ebd.): Noteninflation und befürchteter Qualitätsverlust des Abiturs, Funktion des Abiturs im Hinblick auf den Hochschulzugang, Gestaltung der gymnasialen Oberstufe, Dauer der Schulzeit bis zum Abitur (G8/G9) sowie die Forderung nach verbindlichen und vergleichbaren Prüfungsanforderungen und Rahmenbedingungen (u.a. Zentralabitur, Bildungsstandards). Für vertiefende Feinanalysen wurde zunächst der zuletzt genannte Themenstrang ausgewählt. Der analytische Fokus liegt dabei auf der Rekonstruktion sogenannter „Phänomenstrukturen“ (Keller 2011, S. 248) sowie Deutungsmuster, welche als „typisierende und typisierte Interpretationsschemata“ (ebd., S. 240) gesellschaftliche Wissensbestände organisieren. Die bisherigen Ergebnisse zeigen, dass die AHR in den vergangenen fünfzig Jahren beständig Gegenstand medialer Berichterstattung bzw. öffentlicher Debatten ist (vgl. Autor*in 2022b), wobei im Verlauf des Diskurses jeweils verschiedene Themenstränge zu unterschiedlichen Zeitpunkten präsent sind (vgl. ebd.). Auf Ebene der Deutungsmuster haben sich vor allem die Begriffe „Vergleichbarkeit“, „Qualität“ und „Gerechtigkeit“ als zentral erwiesen, welche im Rahmen der zugrundeliegenden Phänomenstruktur mit unterschiedlichen Deutungsmöglichkeiten versehen sind und den Diskursverlauf entscheidend prägen.

To Score or Not to Score? Machbarkeits- und Performanzfaktoren für automatisches Scoring von Textantworten

Torsten Zesch[1], Andrea Horbach[1], Fabian Zehner[2,3]
[1]FernUniversität Hagen; [2]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; [3]Zentrum für Internationale Bildungsvergleichsstudien (ZIB)

Das Bewerten (d.h., Scoren) von Textantworten in Tests kann inzwischen durch die Verwendung von Methoden zur Verarbeitung natürlicher Sprache und maschinellem Lernen automatisch erfolgen (Yan, Rupp & Foltz, 2020). Ob dies jedoch für ein gegebenes Anwendungsszenario machbar und mit ausreichender Genauigkeit möglich ist, hängt von einer Vielzahl von Faktoren ab. Dieser Beitrag systematisiert diese Faktoren in einem Rahmenmodell auf der Grundlage neuer empirischer Befunde und der Integration empirischer Literaturbefunde, um so Forschenden eine a-priori Abschätzung zu ermöglichen, ob die automatische Auswertung der Textantworten für eine bestimmte Studie oder Testanwendung sinnvoll, effizient und akkurat umsetzbar ist. Wir argumentieren, dass die Performanz (d.h., die Übereinstimmung mit der menschlichen Bewertung) primär von der linguistischen Varianz in den Textantworten abhängt, und erweitern dafür das von Horbach und Zesch (2019) vorgeschlagene Modell um eine zusätzliche Komponente.

Der Vortrag wird die aktuell vorhandenen Methodenparadigmen einführen, die drei Komponenten linguistischer Varianz spezifizieren und das vorgeschlagene Rahmenmodell bezüglich relevanter Einflussfaktoren, die höhere oder niedrigere Varianz bedingen, diskutieren.

Linguistische Varianz und Performanz

Alle Methodenparadigmen haben grundsätzlich gemeinsam, dass sie eher hohe Übereinstimmungswerte mit der menschlichen Bewertung erzielen, wenn die Antworten, die von einer Aufgabe evoziert werden, wenig Varianz aufweisen. Muss das Scoringmodell nämlich nur wenig Informationen über richtige Antworten enthalten, können diese leichter von Expert*innen gebildet oder vom maschinellen Lernen identifiziert werden. Das heißt, wenn eine Aufgabe etwa nur nach dem Namen einer Protagonistin fragt, ist die resultierende Varianz der Textantworten sehr gering, da grundsätzlich nur Personennamen aus dem Stimulus als relevante Antworten in Frage kommen. Horbach und Zesch (2019) sprechen von konzeptueller Varianz, um zu unterscheiden, wie viele semantisch distinkte Konzepte in den Antworten benannt werden. Gibt es im Stimulus etwa nur zwei Personennamen, fällt die Varianz geringer aus, als wenn Tolstois ‘Krieg und Frieden‘ gelesen wurde. Auch der Bewertungsskopus, also ob das Benennen verschiedener Konzepte zur Antwortrichtigkeit führen kann, trägt zur linguistischen Varianz innerhalb der richtigen Antworten bei. Neben richtigen Antworten determinieren außerdem noch die Vielfalt typischer assoziierter Misskonzepte die resultierende sprachliche Varianz in den Antworten.

Ein und dasselbe Konzept kann in natürlicher Sprache häufig auf verschiedenste Weisen ausgedrückt werden ("Philine" vs. "Die Protagonistin heißt Filine"), was von Horbach und Zesch (2019) als Realisierungsvarianz verstanden wird. Bei dieser Komponente kommen sprachliche

Phänomene wie Paraphrasierung oder implizite Sprache zum Tragen, aber auch abweichende Schreibweisen, Grammatik- und Rechtschreibfehler.

Neben der konzeptuellen und Realisierungsvarianz bringt außerdem noch nonkonformes Antwortverhalten Varianz in die Textdaten, etwa durch thematisch abweichende, Nonsense- oder beleidigende Antworten wie auch Täuschungsversuche.

Rahmenmodell für determinierende Faktoren

Während der Abstract die einzelnen Faktoren nur andeuten kann, wird im Vortrag das Zusammenspiel der Faktoren und obengenannten Facetten wie Methodenparadigma und Varianzkomponenten beleuchtet. Das Rahmenmodell unterscheidet Faktoren, die für das automatische System zugänglich sind – wie Antwortlänge, Anzahl der Kodierungskategorien, Trainingsdatenumfang und Testsprache – sowie solche, die für das System nicht sichtbar sind – wie Aufgabentyp, Antwortmodalität, menschliche Übereinstimmungsrate und Populationseffekte. Die Machbarkeit wird vor allem durch die individuelle oder institutionelle Wichtigkeit der Testwertinterpretation determiniert, aber auch rechtlicher und ethischer Rahmenbedingungen, der Anforderung nach individueller Rückmeldung, computationalen Anforderungen und nicht zuletzt finanziellen und ökologischen Kosten.

Conclusio

Der Beitrag systematisiert determinierende Faktoren der Machbarkeit und Genauigkeit automatischer Bewertung. Das vorgeschlagene Rahmenmodell bietet Forschenden die Möglichkeit, im Vorfeld der Studienplanung oder Testanwendung in etwa einzuschätzen, ob die Entwicklung einer automatischen Bewertung machbar und womöglich performant sein wird.

H-08: Wissen und Stress bei Lehrkräften

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S05 T00 B71
Chair der Sitzung: Matthias Krepf

Profile kognitiver Anforderungen zur Beschreibung der unterrichtlichen Planungskompetenz von (angehenden) Lehrpersonen: Eine latente Klassenanalyse

Matthias Krepf, Johannes König
Universität zu Köln, Deutschland

Theoretischer Hintergrund

Die Unterrichtsplanung ist Teil der alltäglichen Arbeit von Lehrpersonen. Obwohl die Unterrichtsplanung eine Kernanforderung im Lehrer*innenberuf darstellt, existieren bisher immer noch wenige empirische Studien, die sich mit der Modellierung und Messung der Unterrichtsplanungskompetenz beschäftigen (Cochran-Smith & Villegas 2016). Ein gemeinsames (Kompetenz-)Konzept lässt sich ebenfalls nicht identifizieren (König et al. 2015). Zahlreiche Arbeiten fokussierten dabei auf das (professionelle) Wissen von Lehrpersonen. In neueren Ansätzen erfolgt eine modellhafte Erweiterung, die stärker *situationsspezifische* Fähigkeiten in den Blick nehmen (Blömeke et al. 2015). Als situationsspezifisch werden *kognitive* Anforderungen der Unterrichtsplanung, die von Lehrpersonen bewältigt werden sollen, angesehen.

Sechs solcher kognitiven Anforderungen wurden im Rahmen des Projekts *PlanvoLL-D* („Die Bedeutung des professionellen Wissens angehender Deutschlehrkräfte für ihre Planung von Unterricht – Planungskompetenz von Lehrerinnen und Lehrern im Fach Deutsch“, König et al. 2020a, b) auf der Basis schriftlicher Unterrichtsplanungen empirisch ermittelt und modelliert.

Im sogenannten CODE-PLAN Modell (*cognitive demands of lesson planning*; König et al. 2021) werden diese kognitiven Anforderungen beschrieben: Transformation von Inhalten (*content transformation*), Erstellung von Aufgaben (*task creation*), Adaptivität an die Voraussetzungen der Lernenden (*adaption of learning dispositions*), Zielklarheit (*clarity of learning objectives*), Kontextualisierung in die Unterrichtseinheit (*unit contextualization*), Phasierung der Unterrichtsstunde (*phasing*).

Diese sechs Anforderungen ließen sich über Teilskalen reliabel messen (α = .55–.83). Mit Blick auf eine Panel-Stichprobe konnte ein statistisch signifikanter und praktisch bedeutsamer Zuwachs der so erfassten Unterrichtsplanungskompetenz von der ersten zur letzten Lehrprobe nachgewiesen werden (*d* = 0.97–1.98 für die sechs Teilskalen). Die Note für die schriftlichen Planungen korrelierte mit der in der letzten Lehrprobe gemessenen Unterrichtsplanungskompetenz signifikant (-.53).

Fragestellung

Das hier vorgestellte CODE-Plan Modell steht exemplarisch für einen *variablenzentrierten* Ansatz. Dieser Ansatz ist am häufigsten in Studien zur Messung der unterrichtlichen Planungskompetenz zu finden. Dadurch lassen sich die durchschnittlichen Ausprägungen von Kompetenzen, Einschätzungen, Motiven etc. angeben.

Deutlich seltener wird der *personenzentrierte* Ansatz gewählt, mit dem die einzelnen (angehenden) Lehrpersonen hinsichtlich ihrer Merkmalsausprägungen in Gruppen mit ähnlichen Ausprägungen zusammengefasst werden. Der variablenzentrierte Ansatz hat den Nachteil, dass bestimmte Merkmalskombinationen nicht erfasst werden. Der personenzentrierte Ansatz schließt diese Lücke

und liefert ein differenzierteres Bild, um die kognitiven Anforderungen akzentuierter beschreiben zu können.

Daher gehen wir den Fragen nach, (1) ob sich die kognitiven Anforderungen mittels latenter Klassenanalysen bündeln lassen und (2) inwieweit die Klassenlösungen zu einer qualitativ (differenzierteren) Beschreibung und somit spezifischeren Modellierung der unterrichtlichen Planungskompetenz beitragen können.

Methode

Zur Bearbeitung der Forschungsfrage greifen wir auf die Ergebnisse der Kodierung des Code-Plan Modells zurück. Dort wurden 310 *schriftliche* Unterrichtsentwürfe von Referendar*innen (172 zu T1 [erste Lehrprobe]; 138 zu T2 [letzte Lehrprobe/Staatsprüfung] aus Nordrhein-Westfalen und Berlin aus dem Projekt PlanvoLL-D (König et al., 2020a, b) ausgewertet. Die Auswertung der Unterrichtsplanungen erfolgt inhaltsanalytisch mittels deduktiv gebildeter Kategorien (Mayring 2014). Die jeweiligen Kodierereignisse wurden anschließend quantifiziert, um mehrere latente Klassenanalysen (LCA) zu berechnen.

Die LCA ist ein personenzentrierter Ansatz der Datenanalyse (Rost, 2004). Sie wird verwendet, um Gruppen oder Subgruppen von Fällen aufzuspüren. Solche Untergruppen werden als „latente Klassen“ bezeichnet. Nicht direkt messbare Konzepte, wie im vorliegenden die Unterrichtsplanungskompetenz, werden über messbare Variablen (sechs kognitiven Anforderungen an Unterrichtsplanung) anhand der Ausprägungen zu unterschiedlichen Klassen gruppiert. Die Klassen geben dann Aufschluss darüber, inwieweit die Anforderungen an die Unterrichtsplanungen umgesetzt wurden und ob sich die Gruppen statistisch signifikant und praktisch bedeutsam voneinander unterscheiden lassen.

Ergebnisse

Die Kodierungen sind abgeschlossen (König et al. 2021). Die latenten Klassenanalysen wurden berechnet und die Anzahl der Klassen über Modellvergleiche bestimmt. Zur Bestimmung der Güte der Klassenlösungen, werden verschiedene Maße herangezogen (Geiser 2011). Mit Blick auf die informationstheoretischen Maße (z.B. AIC, saBIC) und statistischen Tests (z.B. VLMRT) ist eine 3- oder 4-Klassenlösung naheliegend. Aktuell finden die weiterführenden Auswertungen statt. Die endgültigen Ergebnisse liegen zur Tagung vor.

LWK - Lehrerwissen zur Klassenführung erfassen: Konzeption und Validierung eines Wissenstests zur Klassenführung für Lehrkräfte und Lehramtsstudierende der Primarstufe

Elisabeth Seethaler[1], Gerlinde Lenske[2], Bernadette Gold[3], Georg Krammer[4]
[1]Pädagogische Hochschule Salzburg, Österreich; [2]Leuphana Universität Lüneburg; [3]Technische Universität Dortmund; [4]Pädagogische Hochschule Steiermark

Was Lehrkräfte wissen und können sollten, um Expert:in (Berliner, 2001, 2004) im Lehrer:innenberuf zu werden, bildet das Modell *professioneller Handlungskompetenz von Lehrkräften* ab (Baumert & Kunter, 2006). Das darin enthaltene Professionswissen gilt als wesentliche Voraussetzung für erfolgreiches Lehrerhandeln (Kunter et al., 2011) und beinhaltet u. a. das pädagogisch-psychologische Wissen, welches Klassenführungswissen einschließt (Shulman, 1987). Während Studien die allgemeine Relevanz des pädagogisch-psychologischen Wissens nahelegen (z.B. für

Unterrichtsqualität: Lenske, Wirth & Leutner, 2017), gibt es Hinweise darauf, dass die positiven Effekte des pädagogisch-psychologischen Wissens über die Klassenführung vermittelt werden (Lernzuwachs der Schüler:innen: Lenske et al., 2016; verringerte Unterrichtsstörungen: Voss et al., 2014). Bisherige Erfassungen von Klassenführungskompetenzen sind häufig videobasiert und identifizieren somit die professionelle Wahrnehmung (z.B. Gold & Holodynski, 2017). Darüber hinaus ist uns kein Instrument zur Erfassung des Klassenführungswissens bekannt, welches auch die Dimension Beziehung, welche gerade in der Grundschule im Bereich Klassenführung nicht zu unterschätzen ist, berücksichtigt (vgl. Lenske, Bönte, van Bebber & Leutner, 2022).

Aufgrund des Mangels an Paper-Pencil-Tests zur Erfassung des Klassenführungswissens haben wir uns zum Ziel gesetzt, jeweils einen Wissenstest für die Primar- (GS) und die Sekundarstufe (SEK) zu konzipieren und zu validieren. In diesem Beitrag wird die Grundschulversion fokussiert. Dabei werden Testkonzeption bzw. -überarbeitung beschrieben und zentrale Gütekriterien untersucht, wobei die faktorielle Validität, die Änderungssensitivität und die kriteriale Validität besondere Beachtung finden.

Die inhaltliche Testkonzipierung fußt auf dem *Linzer Konzept der Klassenführung* (LKK; Lenske & Mayr, 2015). Aufbauend auf den Vorarbeiten von Kurz und Lenske (2018) wurde zur Test(weiter)entwicklung im ersten Schritt die Methode des kognitiven Pretesting (Beatty & Willis, 2007) mit 56 Lehramtsstudierenden eingesetzt. Die Auswertung erfolgte mit Hilfe der qualitativen Inhaltsanalyse nach Mayring (2010). Im Anschluss wurde ein Abgleich der qualitativen Daten mit den statistischen Kennwerten aus den quantitativen Analysen aus den Vorarbeiten (Trennschärfen, Schwierigkeit, Ladung, vgl. Kurz & Lenske, 2018) vorgenommen. Die Items wurden in einem iterativen Prozess unter Einbindung von Expert:innen gänzlich überarbeitet oder erneuert. Die adaptierte Version umfasst 27 Items, welche entweder im single-choice oder multiple-choice Format vorliegen. Die Items wurden so konzipiert, dass inhaltlich alle Kategorien bzw. Dimensionen des LKK (Beziehung, Kontrolle, Unterricht) mit den dazugehörigen Handlungsstrategien berücksichtigt sind (Lenske & Mayr, 2015). Ferner wurde darauf geachtet sowohl Items zum deklarativen als auch zum konditional-prozeduralen Wissen abzubilden. Die Testdauer beträgt ca. 30 Minuten. Die Variation des Klassenführungswissens in der Zielpopulation bilden wir durch den Unterschied im Schwierigkeitsgrad der Items, der vorab eingeschätzt wurde, ab. Die faktorielle Validität wurde auf Basis von konfirmatorischen Faktorenanalysen durchgeführt. Die Validierung des Tests erfolgt mehrschrittig. Der erste Teil der Validierung (N=304, Lehramtsstudierende aus drei verschiedenen Hochschulen im Bachelor) dient der Überprüfung der faktoriellen Validität und ist bereits abgeschlossen. Weitere Validierungsschritte befassen sich auch mit der Sensitivität des Instruments, um sowohl interindividuelle (Gruppen mit erwartbaren Unterschieden im Klassenführungswissen) und intraindividuelle (erwartbare Unterschiede durch Training von Klassenführung) Unterschiede abzubilden. Auch die kriteriale Validität (Zusammenhänge mit der professionellen Wahrnehmung und der Selbstwirksamkeit) werden überprüft.

Das postulierte dreifaktorielle Modell lässt sich auf Basis der empirischen Daten abbilden ($c^2[374]=386.675$, $p=.315$, RMSEA=.011, CFI=.98, WRMR=0.943), wobei eine Itemselektion angeregt durch Itemkennwerte zu einem parsimonischerem Modell und zu einer Verbesserung des Modellfits führte ($c^2[132]=132.259$, $p=.477$, RMSEA=.003, CFI=1.00, WRMR=0.879). Zusammengenommen

sprechen diese Ergebnisse für die faktorielle Validität des Tests. Kritisch ist anzumerken, dass geringes Vorwissen der Proband:innen bestand, da die Studierenden teilweise gerade das Studium begonnen haben. Es ist daher derzeit noch offen, inwiefern Varianzeinschränkungen durch Bodeneffekte die Validierung beeinflusst haben und wie trennscharf das Instrument bei höheren Ausprägungen des Klassenführungswissens ist. Vor diesem Hintergrund erfolgt bereits der zweite Teil der Validierung im Zuge einer Pre-Post-Testung, wobei die Änderungssensitivität im Zentrum steht. Die Ergebnisse werden im Vortrag ebenfalls berichtet und diskutiert.

Individuelle Stressbewältigungsmuster bei Lehramtsstudierenden? Latente Profilanalysen von Copingstrategien und ihre Korrelate zu Stress und Stressoren

Juliane Schlesier[1], Andrea Westphal[2]
[1]Carl von Ossietzky Universität Oldenburg, Germany; [2]Universität Greifswald, Germany

Theoretischer Hintergrund

„Gelassen durch ungewisse Zeiten“ titelte *Psychologie Heute* im Dezember 2021. Wohl kein anderes Thema dominiert populärwissenschaftliche Psychologiezeitschriften so stark, wie die Frage nach einer angemessenen Bewältigung von Herausforderungen insbesondere während der Corona-Pandemie. Auch die empirische Forschung hat sich mit Copingstrategien von Studierenden befasst (Mahmoud et al., 2012; Matheny et al., 2002; Park et al., 2020; Rogowska et al., 2020) und konnte zeigen, dass Studierende, die funktionale Copingstrategien einsetzen, weniger Stress erleben (Naquin & Gilbert, 1996; Pearman et al., 2021; Rogowska et al., 2020). Doch in den meisten Studien wird nicht berücksichtigt, dass Individuen bestimmte Strategien gleichermaßen präferieren und diese in relativ stabilen Kombinationen einsetzen (Mauno et al., 2014). Zwar gab es in den letzten Jahren vermehrt Bestrebungen, Copingprofile im Rahmen eines personen-zentrierten Ansatzes zu ermitteln (Kavčič et al., 2022; Pété et al., 2022). Es fehlen aber noch immer Untersuchungen zu Stressbewältigungsprofilen von Lehramtsstudierenden, die wichtige Impulse für eine gezielte Förderung eines günstigen Stressbewältigungsverhaltens liefern können.

Fragestellung und Hypothesen

Folglich ist es das Ziel der vorliegenden Studie, mittels einer latenten Profilanalyse typische Stressbewältigungsmuster von Lehramtsstudierenden zu identifizieren. Weiterhin soll untersucht werden, welche Stressbewältigungsprofile besonders förderlich für ein geringes Stresserleben sind. Da bisherige Profilanalysen zu heterogenen Ergebnissen mit einer unterschiedlichen Anzahl an Copingprofilen geführt haben und Untersuchungen mit Lehramtsstudierenden fehlen (Aldridge & Roesch, 2008; Kavčič et al., 2022; Mauno et al., 2014; Pété et al., 2022), ist der Ansatz der vorliegenden Analyse explorativ.

Methode

Stichprobe. Die Untersuchung basierte auf einer Stichprobe von N = 316 Lehramtsstudierenden der Universität Oldenburg, die am Ende des Wintersemesters 2020/21 an einer Online-Befragung teilnahmen. Die Lehramtsstudierenden waren im Durchschnitt 24 Jahre alt (SD = 3.84) und 85.4% der Befragten waren weiblich. Die Befragten waren in unterschiedlichen Lehramtsstudiengängen eingeschrieben.

Instrumente. Erfasst wurden neun Copingstrategien der deutschen Version (Knoll, 2003) des Brief COPE Inventars (Carver, 1997): Selbstablenkung, Verleugnung, emotional-soziale Unterstützung, Aufgeben, Positive Umbewertung, Aktive Bewältigung, Instrumentell-soziale Unterstützung, Ablenkung und Akzeptanz. Alle Copingstrategien (Ausnahme: Selbstablenkung) zeigten akzeptable bis gute interne Konsistenzen (.61 ≤ ω ≤ 86). Weiterhin wurde die deutsche Version (Schneider et al., 2020) der Perceived Stress Scale (PSS; Cohen et al., 1983) eingesetzt, um die Selbstwirksamkeit und die Hilflosigkeit der Studierenden zu ermitteln (ω = .76 und .86). Um spezifische Stressoren im Lehramtsstudium während der Corona-Pandemie zu erfassen, wurde eine selbst entwickelte Skala eingesetzt. Ein exploratives Strukturgleichungsmodell ergab sechs Subskalen (.62 ≤ ω ≤ 86).

Statistische Analysen. Berechnet wurden latente Profilananalysen mit dem Paket tidyLPA (Rosenberg et al., 2018). Den Profilanalysen liegen latente Mittelwerte aus konfirmatorischen Faktorenanalysen der COPE-Skalen zugrunde, die mittels MplusAutomation (Hallquist & Wiley, 2018) in R berechnet wurden. Um die optimale Lösung zu identifizieren, wurden verschiedene Model-Fit-Indices (BIC, SABIC, AIC und CAIC) herangezogen und BLRT, Entropy-Werte sowie die Stichprobenumfänge der Profile berücksichtigt (Spurk et al., 2020; Tein et al., 2013). Basierend auf der finalen Profillösung haben anhand des 3-Schritt-Ansatz (Vermunt, 2010) geprüft, inwiefern sich die Profile in ihrem Stresserleben und der erlebten Stressfaktoren unterscheiden.

Ergebnisse und ihre Bedeutung

Latente Profilanalyse (LPA) ergaben, dass eine Lösung mit sechs Profilen am besten auf die Daten passt. Die sechs identifizierten Profile lassen sich bezeichnen als „Networking“, „Disengaged“, „Withdrawn“, „Cognitive Restructurers“, „Avoidant“ und „Repertoire-oriented Copers“. Individuen, die sich den beiden Coping-Profile zuordnen ließen, die in besonderem Maße auf soziale und emotionale Unterstützung zurückgreifen („Networking Copers“, „Disengaged Copers“), erlebten signifikant weniger Stress. „Avoidant Copers“ und „Withdrawn Copers“ zeigten ein hohes Stresserleben und eine ungünstige Wahrnehmung der Stressoren. Wir diskutieren, wie die Implementation von Programmen zur Förderung der sozial-emotionalen Unterstützung in der Lehramtsausbildung dazu beitragen kann, die Stressbewältigung der Lehramtsstudierenden zu fördern und ihr Stresserleben zu senken.

Wie stabil ist die emotionale Erschöpfung bei angehenden Lehrkräften? Eine Analyse von zwei Längsschnittstudien mittels STARTS-Modellen

Bastian Carstensen[1], Oliver Lüdtke[1,2], Karen Aldrup[1], Uta Klusmann[1]
[1]IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany; [2]ZIB – Zentrum für internationale Bildungsvergleichsstudien

Theoretischer Hintergrund

In Erklärungsansätzen zum beruflichen Beanspruchungserleben wird angenommen, dass Burnout als längerfristige Konsequenz des Ungleichgewichts zwischen beruflichen Ressourcen und Stressoren auftritt (z. B. Bakker et al., 2014). Die Gültigkeit dieser Modelle zeigt sich in empirischen Befunden zu absoluten Niveauveränderungen in der Symptomatik, beispielsweise infolge neuer Herausforderungen nach beruflichen Übergängen (Klusmann et al., 2012). Gleichzeitig deuten hohe Retest-Korrelationen (Hakanen et al., 2011) und Zusammenhänge mit Persönlichkeitsmerkmalen (Roloff et al., 2022; Swider & Zimmermann, 2010) darauf hin, dass einige Personen tendenziell

stärker belastet sind als andere. Eine hohe Rangordnungsstabilität im Ausmaß von Burnout zeigte sich beispielsweise auch für Lehrkräfte im Vorbereitungsdienst (Zimmermann et al., 2016). Da Retest-Korrelationen jedoch durch Messfehler und die Länge des betrachteten Zeitintervalls beeinflusst werden (Anusic & Schimmack, 2016), lässt sich aus diesen nicht unmittelbar ableiten, in welchem Ausmaß die Stabilität und Variation interindividueller Unterschiede durch absolut stabile Ursachen, langfristige Prozesse oder zeitlich begrenzte Faktoren verursacht werden. Ein tieferes Wissen über die Ursachen der interindividuellen Stabilität von Burnout-Symptomen könnte das theoretische Verständnis über zugrundeliegende Entstehungs- und Veränderungsprozesse erweitern und praktische Fragen zur Prävention und Intervention klären.

Eine Möglichkeit zur Analyse der Stabilität von Messungen bieten STARTS-Modelle (Kenny & Zautra, 2001), die interindividuelle Unterschiede über die Zeit in drei Anteile zerlegen: (1) eine zeitlich stabile Komponente, die auf stabilen Ursachen wie der individuellen genetischen Disposition beruht; (2) eine zeitlich variierende, teilweise stabile Komponente, die vornehmlich durch systematische Prozesse über die Zeit zustande kommt; sowie (3) eine komplett situationsspezifische Komponente, die messzeitpunktspezifische Faktoren sowie Messfehler repräsentiert. Das STARTS-Modell wurde aufgrund hoher Anforderungen an die Daten und häufiger Schätzprobleme bisher nur selten in der Forschungspraxis eingesetzt. Für die emotionale Erschöpfung, das Leitsymptom des Burnout-Syndroms, konnten Dicke et al. (2021) in einer Anwendung dieses Modells aufzeigen, dass sie sich bei Schulleiter*innen zu etwa gleichen Teilen aus stabilen, autoregressiven und situationsspezifischen Komponenten zusammensetzte. Basierend auf diesen Erkenntnissen zielte die vorliegende Studie darauf ab, die Zusammensetzung emotionaler Erschöpfung bei Lehramtsstudierenden zu untersuchen und der Frage nachzugehen, wie stabil die emotionale Erschöpfung in dieser Population ist.

Methode

Unsere Fragestellungen haben wir mit Angaben von N = 4,045 Lehramtsstudierenden aus der *Student Teacher Professional Development Study* (*STePS*, Carstensen et al., 2021) sowie Daten von N = 2,034 Lehramtsstudierenden des *Panels zum Lehramtsstudium* (*PaLea*, Bauer et al., 2010) untersucht, wobei jeweils ein Zeitraum von insgesamt fünf Jahren betrachtet wurde. Erfasst wurde die emotionale Erschöpfung in beiden Panelstudien jährlich mittels adaptierter Versionen des Maslach Burnout Inventory (Enzmann & Kleiber, 1989). Die interne Konsistenz fiel in beiden Studien für alle Messzeitpunkte zufriedenstellend aus ($.80 < \alpha < .91$). Wir nutzten das R-Paket STARTS (Robitzsch & Lüdtke, 2018) zur Analyse der Varianzanteile emotionaler Erschöpfung.

Ergebnisse und Diskussion

Die emotionale Erschöpfung wies in beiden Stichproben substanzielle Varianz in allen Komponenten auf. Gemittelt über beide Stichproben lag der stabile Anteil im untersuchten 5-Jahres-Zeitraum bei ca. 36%, wohingegen die situationsspezifische Komponente im Mittel ca. 27% der Gesamtvariation ausmachte. Schließlich hatte die autoregressive Komponente mit ca. 41% den größten Varianzanteil in der emotionalen Erschöpfung. Durch die eher hohe autoregressive Stabilität von $\beta \approx .71$ ergibt sich

für die modellgeschätzte längsschnittliche Korrelation eine relativ langsame Verminderung über die Zeit, bis diese sich langfristig dem Wert des stabilen Anteils asymptotisch annähert.

Insgesamt deckt sich das Ergebnismuster mit den Annahmen der etablierten Modelle zur Stressgenese, die vor allem die Interaktion zwischen dem Individuum und der Umwelt als bestimmenden Faktor für die Manifestation von Burnout-Symptomen diskutieren (Bakker et al., 2014; Lazarus & Folkman, 1984). Entsprechend empfehlen sich in praktischer Hinsicht insbesondere Präventions- und Interventionsmaßnahmen, die an diesen Interaktionsprozessen ansetzen. Gleichzeitig sollten stabile Personenfaktoren als Ursache für die relative Stabilität emotionaler Erschöpfung in zukünftiger Forschung näher untersucht werden.

H-09: Cheating und automatisches Scoring

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S05 T00 B83
Chair der Sitzung: Martin Daumiller

„Nicht für mich, sondern für andere“: Anderen bei akademischem Betrugsverhalten zu helfen hängt mit Beziehungszielen und kooperativen sozialen Normen zusammen

Martin Daumiller[1], Tanja Fritz[1], Hernán González Cruz[2], Selma Rudert[3], Stefan Janke[2]
[1]Universität Augsburg, Germany; [2]Universität Mannheim, Germany; [3]Universität Koblenz-Landau, Germany

Theoretischer Hintergrund

Akademisches Betrugsverhalten hat negative Auswirkungen auf die Fairness innerhalb von Bildungssystemen. Neben aktivem Betrügen umfasst dieses auch Beihilfe hierzu (z.B. andere von einem abschreiben lassen), was ähnlich schädlich und prävalent, allerdings bisher kaum untersucht ist (Putarek & Pavlin-Bernardic, 2020).

Prädiktoren entsprechenden Verhaltens nachzugehen ist somit eine wichtige Forschungsbemühung. Bedeutsam erscheinen hierzu, wie bei aktivem Betrugsverhalten, neben institutionellen Vorgaben Ziele und wahrgenommene Normen der Studierenden (Daumiller & Janke, 2020): Konkret sollten selbstbezogene Ziele (zu deren Erreichung Betrugsverhalten zu-/abträglich sein kann) sich auch auf die Bereitschaft zu Betrugs-Beihilfe auswirken: Weniger relevant scheinen hier allerdings die bisher in der Forschungsliteratur betrachteten und in ersten Studien nicht signifikanten Master- und Performanzziele (Pavlin-Bernardic et al., 2017). Vielmehr postulieren wir, dass *Beziehungsziele*, die auf zwischenmenschliche Beziehungen fokussieren (hier: mit Kommiliton*innen), Beihilfe zum Betrug bahnen. So kann Betrugsbeihilfe ein Mittel darstellen, um sozialen Anschluss zu finden/halten (siehe Bucciol et al., 2020).

Für die Wirkung der Ziele erscheinen zudem soziale Normen relevant. Insbesondere wenn in einem Umfeld in besonderem Ausmaß kooperatives Verhalten wertgeschätzt wird (starke *Kooperationsnorm*; Wagner, 1995), ist davon auszugehen, dass Studierende auch bei Prüfungen eher beim Betrügen unterstützen. Gleichzeitig ist auch davon auszugehen, dass Betrugsbeihilfe auch als hilfreiches Mittel zum Aufbau/Erhalt kooperativer Beziehungen angesehen wird. Dies bedeutet, dass Kooperationsnormen sowohl Haupteffekte auf Betrugsbeihilfe entfalten als auch die Effekte von Beziehungszielen moderieren sollten.

Fragestellung

Hinsichtlich einer bislang wenig betrachteten, schädlichen und prävalenten Form akademischen Betrugsverhaltens, für die andere motivationale Mechanismen vorzuliegen scheinen, untersuchten wir zwei Vorläuferbedingungen. Wir erwarteten, dass Beziehungsziele positiv mit Beihilfe zu akademischem Betrugsverhalten zusammenhängen und testeten kooperative sozialer Normen als weiteren Prädiktor und Moderator.

Methode

Datengrundlage unserer präregistrierten, deutschlandweiten Studie waren 835 Studierende (davon 607 weiblich; durchschnittliches Alter: 23.2, *SD*=3.6, Jahre), die im Sommersemester 2020 Ziele und Normen und im nachfolgenden Semester, auf den Prüfungszeitraum rückblickend, Betrugsverhalten berichteten.

Beziehungsziele wurden mit der Studierendenversion der Skala von Daumiller et al. (2019) gemessen (4 Items, gute interne Konsistenz: McDonald's ω_h=.92). *Kooperative soziale Normen* wurden in Hinblick auf jene Komiliton*innen mit denen man viel Zeit im Studium verbringt mit fünf Items von Chatman et al. (2001) erhoben (ω_h=.86).

Beihilfe zu Betrugsverhalten erfassten wir mit drei adaptierten Items (Genereux & McLeod, 1995; Schmelkin et al., 2010) zu: Weitergabe von Prüfungsaufgaben, Kollaboration mit anderen bei Einzelarbeitsaufgaben sowie Ableisten von Prüfungen/Hausarbeiten für andere. Diese wurden zur besseren Beantwortbarkeit auf einer Häufigkeitsskala von *nie* bis *sehr häufig* präsentiert („nie" wurde jeweils von 47%, 43% und 91% der Befragten angegeben), anschließend dichotomisiert und addiert. Dieser Summenscore repräsentiert hinsichtlich wie viele der drei Aspekte die Befragten Beihilfe zu Betrugsverhalten gezeigt hatten (Janke et al., 2022).

Ergebnisse und Bedeutung

Knapp ein Drittel berichtete keinerlei der drei Verhaltensweisen, 28% eine, 38% zwei und 6% alle drei, wobei 6% angaben, schon einmal eine Prüfung für andere geschrieben zu haben, also zu substanziellem Ausmaß Betrugsbeihilfe geleistet zu haben. Strukturgleichungsmodelle zeigten, dass sowohl Beziehungsziele (β=.08, *p*=.03) als auch kooperative soziale Normen (β=.17, *p*=.001) Betrugs-Beihilfe prädizierten. Ihre Interaktion war nicht signifikant und die Befunde robust für den Umfang an Präsenz/Online-Studium.

Zur Absicherung der Spezifität von Betrugsbeihilfe betrachteten wir zudem dessen Zusammenhänge mit eigenem Abschreiben in Prüfungen und Plagiarismus (geringe Korrelationen von *r*=.17–.22). Zudem bestätigten wir, dass Lern-, Annäherungsperformanz-, Vermeidungsperformanz- und Arbeitsvermeidungsziele nicht signifikant mit Betrugsbeihilfe zusammenhingen.

Zusammengenommen zeigt dies, wie soziale Motivationen akademisches Betrugsverhalten bahnen können. Gerade in einem starken sozialen Kontext können Beziehungsziele und kollektive soziale Normen eine besondere Gefahr für akademische Integrität darstellen. Für die Praxis impliziert dies eine stärkere Aufklärungsarbeit über die Schäden von Betrugsbeihilfe und eine Berücksichtigung der starken sozialen Handlungsgründe hinter diesem Verhalten. Hierfür relevante Kontextbedingungen zu identifizieren kann demnach als besonders wichtig erachtet werden, um deren ungünstige Effekte auf Betrugsbeihilfe einzudämmen.

Do students come clean after the fact? Leveraging the Benford Randomized Response Technique to more accurately predict student cheating behavior during exams

Hernán González Cruz[1], Tanja Fritz[2], Martin Daumiller[2], Stefan Janke[1]
[1]University of Mannheim, Germany; [2]University of Augsburg, Germany

Cheating during university exams is a common phenomenon that is detrimental to the accurate and fair assessment of academic performance in higher education (Surahman & Wang, 2022). Given its severe implications, it is important to come to a clear understanding of the prevalence of cheating behavior. However, the accurate estimation of cheating prevalence is challenging when using self-report measures due to social desirability. The Benford Randomized Response Technique (Diekmann, 2012) is a method that was specifically developed to address this challenge and has the

distinct advantage of minimizing the amount of noise that must be introduced in the data while preserving participant anonymity. To our knowledge, this method has not yet been applied to measure cheating in the context of a specific exam, an area in which it is particularly difficult to ensure participants' anonymity. Thus, we apply Benford RRT in a student sample to assess cheating and its associations with the three central aspects of Murdock and Anderman's (2006) theoretical framework for dishonest behavior (i.e., motivation, perceived capabilities, and costs) in the context of a specific exam. Specifically, we measured students' seminar-specific learning/performance goals and exam-specific self-efficacy as key aspects of motivation and perceived capabilities. We also measured general attitudes towards cheating, examination format (in-person, online "live", and online "take-home"), and perceived peer behavior as key aspects of personal costs.

We conducted a cross-sectional survey study and questioned the participating university students directly after a summer examination period on their cheating behavior (n = 338 participants, average age = 22.02 years, 64.79% female). We included two retrospective measures of collaborative and illicit cheating during a single exam with an adapted version of the Benford RRT (Diekmann, 2012; Mangat, 1994), as well as measures for the aforementioned predictors (i.e., seminar-specific student goals, exam-specific self-efficacy, general attitudes towards cheating, examination format, and peer behavior).

To analyze our data, we apply logistic regression with mixed effects appropriate for nested RRT data (Heck & Moshagen, 2018). Prevalence estimation revealed that about 21% of participants in our sample cheated during their exam (95% CI [.17, .26]). Furthermore, we found that positive attitudes towards cheating (OR = 2.26, p = .002) and reported cheating behavior of peers (OR = 1.61, p = .03) were positively associated with cheating. A linear contrast for examination format also revealed a significant effect of the online "live" examination format as compared to the online "take-home" examination format (OR = .11, p = .01). Prevalence estimates for the different examination formats suggest that cheating was in fact more common in the "live" format (RRT prevalence estimate = .31, 95% CI [.22, .41]) as compared to the "take-home" format (RRT prevalence estimate = .17, 95% CI [.12, .22]).

Our findings are in line with previous literature and underscore the importance of attitudes and peer behavior as consistently strong predictors of cheating. Differences in online examination formats also suggest that strategic examination planning to reduce the value of cheating for the individual (e.g., process- rather than evaluation focus) may complement efforts to implement rigorous proctoring in online settings, an approach that may be associated with more frequent cheating (Whisenhunt et al., 2022). A substantial prevalence estimate of cheating in our data indicates that, in general, participants answered honestly. Thus, our findings support the use of Benford RRT as a powerful tool to address research questions that require more accurate measurement of sensitive outcomes by ensuring participants' anonymity. This technique can be easily expanded to encompass additional time points as well as additional level-2 predictors (e.g., such as the instructor's goals), which increases its potential use case in educational settings.

Herausforderungen in der experimentellen Untersuchung von Betrugsverhalten: Ergebnisse eines Experiments zur Interaktion von Performanzzielen und Art der Leistungsbeurteilung

Tanja Fritz[1], Hernán Gonzáles Cruz[2], Stefan Janke[2], Martin Daumiller[1]
[1]Universität Augsburg, Germany; [2]Universität Mannheim, Deutschland

Theoretischer Hintergrund. Persönliche Ziele in Leistungssituationen gelten als motivationale Einflussfaktoren von akademischem Betrugsverhalten (Anderman & Koenka, 2017). Theoretisch wird dabei ein positiver Zusammenhang zwischen Performanzzielen und akademischem Betrugsverhalten postuliert. Da diese Ziele den Wert der Leistungsdemonstration in den Vordergrund stellen, können unlautere Mittel eine effektive Strategie zur Zielerreichung darstellen. In der Forschungsliteratur finden sich allerdings sehr heterogene Zusammenhangsmuster zwischen Performanzzielen und akademischem Betrugsverhalten (Daumiller & Janke, 2019; Krou et al., 2021), was auf weitere relevante Faktoren hinweist. Gemäß der Erwartungs-Wert Theorie der Handlungsmotivation (Wigfield & Eccles, 2000) lässt sich davon ausgehen, dass Performanzziele (als Aspekt des Wertes aufzufassen) insbesondere dann Betrugsverhalten bahnen, wenn Individuen zu der Erwartung kommen, dass dieses Verhalten a) zielführend ist und b) eher nicht entdeckt wird (starke Erfolgserwartung in Bezug auf Betrugsverhalten). Ob Studierende entsprechende Schlüsse ziehen, hängt dabei möglicherweise von der Aufgabenart ab. So ist anzunehmen, dass Performanzziele besonders dann zu mehr Betrugsverhalten führen, je stärker die korrekte Lösung (Ergebnisfokus = Betrug eher einfach, niedrige Chance entdeckt zu werden) als der Lösungsprozess (Prozessfokus = Betrug eher schwierig, erhöhte Chance entdeckt zu werden) beurteilt/belohnt wird.

Methode. Um diese Annahme zu überprüfen implementierten wir ein 2 (Performanzzielinduktion vs. keine Induktion) x2 (ergebnis- vs. prozessorientierte Leistungsbewertung)- Between-Subjects Experiment, an dem 238 Studierende (Semesteranzahl: *M* = 4.5, *SD* = 1.8, 67.0% weiblich, 1.3% divers) teilnahmen. Das Untersuchungsdesign und die Analysemethodik wurden vorab präregistriert (https://aspredicted.org/CPJ_HX4) und die erforderliche Stichprobengröße durch eine a priori Power-Analyse bestimmt. Die Versuchsteilnehmenden bearbeiteten in Gruppen (*M* = 2.4, *SD* = 1.0) einen vorgeblichen akademischen Berufseignungstest mit zwei Aufgabentypen (Wörter aus Anagrammen bilden und Wissensfragen). Performanzziele wurden induziert, indem die Teilnehmenden angewiesen wurden einen besonders guten Eindruck zu machen. Zudem wurde ihnen eine zusätzliche Vergütung für gute Leistung in Aussicht gestellt. Gruppen ohne Zielinduktion erhielten keine solcher Hinweise. Der Beurteilungsfokus wurde durch die Aufgabengestaltung selbst sowie Hinweise zur Beurteilung manipuliert: In den Prozessbedingungen hatten die Teilnehmenden die Möglichkeit, ihren Denkprozess zur Lösungsfindung zu notieren und wurden darauf hingewiesen, dass dieser für die Bewertung maßgeblich ist. In den Ergebnisbedingungen wurde den Teilnehmenden erklärt, dass ausschließlich ein richtiges Ergebnis zählt. Zwei Formen von Betrugsverhalten wurden gemessen: wenn mindestens eines von drei tatsächlich unlösbaren Anagrammen als gelöst angegeben wurde, sowie beobachtetes Betrugsverhalten in Form von unerlaubten Hilfsmitteln (z.B. Internet-Recherche).

Ergebnisse. Es zeigten sich keine signifikanten Interaktionseffekte der beiden unabhängigen Variablen, sowie kein Haupteffekt der Performanzzielinduktion. Entsprechend der Annahmen gaben Teilnehmende zwar in der Bedingung mit Zielinduktion und Ergebnisfokus am häufigsten unlösbare Anagramme als gelöst an (50.9% vs. 30.4%–39.4%), was inferenzstatistisch aber nur auf einen

Haupteffekt des Beurteilungsfokus zurückzuführen war (44.6% Ergebnisfokus-Gruppen vs. 31.6% Prozessfokus-Gruppen, β = 0.58, *p* = .017, OR = 1.78). Für beobachtetes Betrugsverhalten waren weder die Zielinduktion, noch der Beurteilungsfokus signifikante Prädiktoren, wobei unerwarteterweise die höchste Betrugsrate unter Zielinduktion und Prozessfokus gemessen wurde (34.6% vs. 14.6%–16.9%).

Diskussion. Während die erwarteten Interaktionseffekte nicht nachgewiesen werden konnten, erhöhte ein ergebnisbezogener Bewertungsfokus das Ausmaß an Betrugsverhalten zumindest bei einem Teil der eingesetzten Aufgaben. Ein auf Verständnis abzielender Bewertungsfokus scheint somit ein effektiver Ansatzpunkt um Betrugsverhalten zu minimieren, auch da dieser Aspekt von Lehrpersonal einfacher zu beeinflussen ist. Wir konnten keinen Effekt von Performanzzielinduktion auf Betrugsverhalten nachweisen. Eine an den Test anschließende Befragung ergab, dass die geringe persönliche Wichtigkeit des Testergebnisses ein häufiger Grund für den Unterlass für Betrug war. Dies lässt die Schlussfolgerung zu, dass eine geringe persönliche Bedeutsamkeit eines Testergebnisses den Effekt von Zielinduktion vermindern kann. Aus einer methodischen Perspektive bedeutet dies, dass zukünftige experimentelle Forschung in noch stärkerem Ausmaß die persönliche Bedeutsamkeit guter Leistung adressieren sollte, um mögliche Zieleffekte (und deren Interaktion) sichtbarer zu machen.

Mehr als nur Technologie- und Fairnessfrage: Ethische Prinzipien beim automatischen Bewerten von Textantworten aus Tests

Fabian Zehner[1,2], Torsten Zesch[3], Andrea Horbach[3]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; [2]Zentrum für Internationale Bildungsvergleichsstudien (ZIB); [3]FernUniversität Hagen

In Anlehnung an etablierte Testgütekriterien elaboriert dieser Beitrag ein Rahmenwerk zur ethischen Bewertung automatischer Auswertung von Textantworten.

Der Einsatz automatischer Bewertung von Textantworten in Tests durch Methoden der natürlichen Sprachverarbeitung (NLP) und des maschinellen Lernens hat sich über die letzten zwei Jahrzehnte sukzessive verbreitet (Yan, Rupp & Foltz, 2020). Obgleich dieser Fortschritt aufgrund diverser Vorteile gegenüber der menschlichen Beurteilung grundsätzlich wünschenswert ist, müssen sowohl Technologieentwicklung als auch -einsatz im Rahmen ethischer Prinzipien kritisch hinterfragt werden. Nationale (DIN 33430:2016-07, 2016) wie auch internationale (ISO 10667-1:2011-10, 2011) Teststandards setzen ohnehin einen klaren Rahmen, in denen sich Testanwendungen und deren Auswertung bewegen dürfen (AERA/APA/NCME, 2014), der Einsatz künstlicher Intelligenz wird auf Ebene der Europäischen Union reguliert (European Commission, 2020) und auch die American Psychological Association (APA) stellt zur Automatisierung von Assessments Folgendes klar: (i) “[Psychologists] select scoring and interpretation services (including automated services) on the basis of evidence of the validity” und (ii) “[they] retain responsibility for the appropriate application, interpretation, and use of assessment instruments, whether they score and interpret such tests themselves or use automated or other services” (APA, 2017, p. 14). Dementsprechend argumentieren wir in diesem Beitrag, dass es eines Regelwerks für den Einsatz automatischer Testauswertung bedarf, das über Fairness hinaus geht. Dafür strukturieren wir die Anwendung ethischer Prinzipien

gemäß etablierten Testgütekriterien (AERA/APA/NCME, 2014, ergänzt durch Kubinger, 2019) im Lichte des Evidence-Centered Designs (Mislevy, Almond & Lukas, 2003) und ethischen Traditionen im Bereich der künstlichen Intelligenz sowie gesellschaftlichen Werten als leitenden Prinzipien.

Testgütekriterien, Werte und automatische Auswertung. Werte stellen leitende Prinzipien dar, über die sich eine Gemeinschaft mehr oder wenig einig ist. Die APA formuliert fünf Prinzipien (APA, 2017): Beneficence and Nonmaleficence, Fidelity and Responsibility, Integrity, Justice und Respect for People's Rights and Dignity. Im NLP-Kontext stellt außerdem der Datenschutz einen genuinen Wert dar. Während ethische Bewertungen von Assessments stets von diesen Werten geleitet sein müssen, können sie ganz praktisch entlang der Taxonomie von Testgütekriterien strukturiert werden. Diese werden im Folgenden ausschnittsweise aufgezeigt, in der Präsentation aber tiefer elaboriert.

Allem voran muss durch das Zeigen von *Fairness* sichergestellt werden, dass automatische Bewertungen menschlichen Bias nicht verstärken (Bridgeman, Trapani & Attali, 2009), sondern vielmehr durch methodische Maßnahmen reduziert werden (Loukina, Madnani & Zechner, 2019). NLP hat darüber hinaus eigene Bias-Quellen, wie etwa die Anfälligkeit für die inakkurate Verarbeitung nicht-konventionellen Sprachgebrauchs (etwa bei weniger fähigen Populationen).

In Hinblick auf die *Validität* von Testwertinterpretationen bringt NLP üblicherweise den Vorteil mit sich, mit offenem Antwortformat die Konstruktabdeckung, bzw. Inhaltsvalidität, verbessern zu können. Andererseits könnte NLP relevante Antwortmerkmale ignorieren und so die Inhaltsvalidität gefährden. Entsprechend dem Evidence-Centered Design (Mislevy et al., 2003) muss außerdem für das Erstellen einer Diagnose vergleichbare Evidenz evoziert und beobachtet werden. Daraus ergibt sich, dass für valide Testwertinterpretationen automatische Testauswertung nicht auf beliebige Textprodukte von Personen angewendet werden kann. So wird etwa illegitimerweise regelmäßig behauptet, Intelligenz anhand von beliebigen Texten messen zu können (z. B. Abramov & Yampolskiy, 2019; McDaniel, 2006).

Durch konsistente Bewertungen kann automatische Auswertung die *Objektivität* einer Messung üblicherweise verbessern.

Reliabilität der Testwerte muss unverändert gezeigt werden, allerdings kann NLP die Messeffizienz erhöhen, indem sie computer-adaptives Testen bei offenem Antwortformat ermöglicht.

Weitere besonders zu berücksichtigende Gütekriterien stellen die *Ökonomie*, *Nützlichkeit* und *Akzeptanz* sowie *Unverfälschbarkeit* dar. Außerdem muss der Skopus der jeweiligen Anwendung berücksichtigt werden: Wie viele Personen sind von der Testung und ihren Konsequenzen betroffen? Wie einschneidend ist der Einfluss für die Betroffenen? Betreffen Konsequenzen der Testwertinterpretation Individuen oder Systeme?

Anwendungsbeispiel und Conclusio. In der Präsentation wird ein Anwendungsbeispiel einer ethischer Bewertung mithilfe des beschriebenen Rahmenwerks demonstriert. Der Beitrag zeigt die Relevanz der ethischen Betrachtung der Testwertverwendung bei automatischer Auswertung auf und bietet ein praktisches Rahmenwerk.

H-10: Lernen in MINT-Fächern

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S05 R03 H20
Chair der Sitzung: Anja Friedrich

Die Rolle des MINT-Unterrichts beim Aufbau von Data Literacy: Ein systematisches Review

Anja Friedrich[1], Saskia Schreiter[2], Sarah Malone[1], Stephanie Haaß[2], Sebastian Becker-Genschow[3], Jessica Lehmann[1], Tobias Mosetter[2], Jochen Kuhn[4], Markus Vogel[2], Roland Brünken[1]
[1]Universität des Saarlandes, Deutschland; [2]Pädagogische Hochschule Heidelberg; [3]Universität zu Köln; [4]Ludwig-Maximilians-Universität München

Der kompetente Umgang mit Daten und ihren Darstellungsformen – genannt *Data oder Statistical Literacy (DSL)* – zählt zu den Schlüsselkompetenzen des 21. Jahrhunderts (z.B. Chalkiadaki, 2018). Gerade in der aktuellen Pandemielage wird deutlich, wie notwendig es ist, dass sich Bürger*innen kritisch mit Darstellungen von Daten und ihren Interpretationen auseinandersetzen können (Watson, 2006; Gould, 2007; OECD, 2021). "Statistical Literacy" wird definiert als die Kompetenz korrekte Schlussfolgerungen aus statistischen Daten zu ziehen, diese kritisch zu beurteilen und darüber zu kommunizieren (Wallman, 1993; Gal, 2002). Gould (2017) erweiterte diese Definition um Datenerfassung und Erstellung von Datendarstellungen zu "Data Literacy". Konsens im Hinblick auf diese Definitionen lässt sich jedoch in einschlägigen empirischen Arbeiten nicht finden, weder im Hinblick auf die Trennung der Begriffe noch bzgl. deren Überschneidung bzw. Erweiterung.

Nicht nur die Definitionen der Begriffe haben sich in den letzten Jahren erweitert, sondern auch deren Bedeutung in Gesellschaft und Forschung (OECD, 2001). Hieraus ergibt sich die Anforderung eben diese Kompetenzen effektiv und zukunftsorientiert zu fördern. Vorangegangene Übersichtsarbeiten haben vielversprechende Methoden zur Förderung von DSL identifiziert, die sich auf Erwachsene und den Hochschulbereich beziehen (Gal, 2002; Schüller et al., 2019). Neuere Forschungsarbeiten nehmen auch DSL im Kontext der schulischen MINT-Bildung in den Blick, da dort verwandte Grundkonzepte Eingang in Lehrpläne und Bildungsstandards gefunden haben (z.B. NCTM, 2000; Kultusministerkonferenz, 2020). Um diese Ansätze weiterzuentwickeln und einen konzeptionellen Rahmen für DSL im Kontext des MINT-Unterrichts zu schaffen, wurden für das hier durchgeführte systematische Review folgende Fragestellungen formuliert:

1. Welche Definitionen von DSL werden in der Forschung im Kontext des MINT-Unterrichts verwendet und konzeptualisiert?
2. Welche Ansätze zur Förderung von DSL im MINT-Unterricht werden empirisch untersucht?
3. Welche Schülervariablen werden in Studien zu DSL im MINT-Unterricht erfasst?

Basierend auf den aktuellen PRISMA-Richtlinien (Page et al., 2021) wurde zunächst eine breite systematische Suche in PsychInfo, ERIC und Web of Science nach internationalen Originalstudien durchgeführt, die explizit DSL im schulischen MINT-Kontext untersuchten. Die systematische Suche resultierte in 16.791 Treffern. Mit Unterstützung des KI-basierten Programms ASReview (Utrecht University, 2021), das eine dynamische Sortierung der gefundenen Publikationen vornimmt, wurden in einem ersten Titel-Abstract-Screening 594 Publikationen identifiziert, die möglicherweise den Inklusionskriterien dieses Reviews entsprechen. Von diesen wurden im Volltexte-Screening 62 Publikationen für die finale Datengrundlage identifiziert. Beide Screenings wurden von zwei Personen (Interrater-Reliabilitäten: Cohen's Kappa=.568/.888 für Titel-Abstract-Screening/Volltextscreening) mit den folgenden Auswahlkriterien durchgeführt:

- Abstract in Englisch

- Originäre Forschung

- Forschungsgegenstand bezieht sich auf K-12 MINT Bildung

- Schüler*innen-Variablen werden berichtet

- Peer reviewed

- Volltext beinhaltet die Begriffe data litera* or statistical litera*

Zusätzlich wurden durch eine erweiterte Suche in internationalen Konferenzbeiträgen (Jahre 2019-2021) 3 weitere Studien in die Analyse einbezogen. Die relevanten Informationen der finalen Publikationen wurden anhand eines Kodierungsschemas extrahiert und die Ergebnisse mittels qualitativer Inhaltsanalyse (Software MAXQDA) synthetisiert.

Bezogen auf die Frage nach den Definitionen und der Konzeptualisierung von DSL (FS 1) zeigt sich, dass der Begriff „Statistical Literacy“ im Vergleich zu „Data Literacy“ innerhalb der Studien im Mittelpunkt steht und dass überwiegend Definitionen bekannter Autoren (z.B. Wallman, 1993; Watson, 2006, Gal, 2002) Verwendung finden. Die Studien legen ebenfalls nahe, dass dem DSL eine hierarchische Struktur mit verschiedenen Leistungsstufen zugrunde liegen könnte. Circa die Hälfte der Studien konzentriert sich auf spezifische statistische Konzepte, wie z.B. Datenverteilung (Variation, Streuung) (FS 3). Verschiedene effektive Ansätze zur Förderung von DSL lassen sich identifizieren (FS 2). Diese sind vorwiegend in der Mathematik oder fächerübergreifend angesiedelt und integrieren meist problem- und projektorientierten Ansätze, die den Einsatz authentischer Problemstellungen und „realer“ Daten fokussieren. Die Ergebnisse sollen zu einem konzeptionellen Rahmen für die Förderung von DSL im MINT-Unterricht beitragen und insbesondere Forschungslücken und theoretische Fragen aufzeigen, die in zukünftiger Forschung betrachtet werden könnten.

To choose or not to choose? Interessengestützte Differenzierung im Fach Chemie

Fabien Güth, Helena van Vorst
Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund

Die Auswahl von Lernmaterialien mit unterschiedlichen Themen wird in der Literatur als interessenbasierter Differenzierungsansatz beschrieben, jedoch fehlt es an empirischen Ergebnissen hinsichtlich lernförderlicher Effekte (Tomlinson et al., 2003). In den Naturwissenschaften deuten Forschungsergebnisse zum kontextorientierten Lernen an, dass sich unterschiedliche lebensweltliche Anwendungsbezüge – so genannte Kontexte – zur Differenzierung anbieten (Autor & Autor, im Druck). Auf Grundlage der bisherigen Ergebnisse bleibt jedoch unklar, ob die Lernenden von selbstgewählten Kontextaufgaben profitieren.

In der Selbstbestimmungstheorie von Ryan und Deci (2017) wird die Bereitstellung von Wahloptionen als Möglichkeit zur Autonomieunterstützung der Lernenden verstanden. Evidenzen verdeutlichen jedoch, dass die Bereitstellung von Wahloptionen nicht per se lernförderlich ist, sondern eine Passung der Wahloptionen zu den Interessen und Werten der Lernenden hergestellt werden muss (Flowerday & Shell, 2015; Patall et al., 2008). Zugleich bleibt unklar, inwieweit die Ergebnisse auf

kontextorientiertes Lernen in den Naturwissenschaften übertragbar sind. Aus diesem Grund widmet sich die vorliegende Studie folgender Forschungsfrage: Welchen Einfluss hat die Passung zwischen Kontextaufgabe und Lernenden sowie die selbstständige Kontextwahl auf den Lernzuwachs, das situationale Interesse und die kognitive Belastung?

Methodik

Es wurde eine experimentellen Studie im Prä-Post-Design mit drei Treatmentgruppen durchgeführt:

(1) Die Lernenden wählen eine Kontextaufgabe selbstständig (choose & match).

(2) Die Lernenden bekommen eine passende Kontextaufgabe zugewiesen (no choose & match).

(3) Die Lernenden bekommen eine nicht passende Kontextaufgabe zugewiesen (no choose & no match).

Die Zuweisung der Kontextaufgaben für die zweite und dritte Treatmentgruppe wurde mithilfe eines Vorhersagemodell durchgeführt, welches durch maschinelles Lernen entwickelt wurde (Kuhn & Silge, 2022). So wurde auf Grundlage vorab erhobener Personenmerkmale eine kontextbasierte Aufgabe automatisiert vorgeschlagen.

Die Untersuchung wurde mit 219 Lernenden des dritten Lernjahres an sieben Gesamtschulen und Gymnasien im regulären Chemieunterricht durchgeführt. Im Prätest wurden u.a. das Vorwissen in Chemie (Celik, im Druck), das Interesse an Chemie (u.a. Wild & Krapp, 1995) und das chemiebezogene Selbstkonzept (Hoffmann et al., 1998) als Kontrollvariablen erfasst, die zugleich zur Vorhersage einer Kontextaufgabe für die zweite und dritte Treatmentgruppe genutzt wurden. Ferner wurde das Fachwissen im Inhaltsfeld „saure und alkalische Lösung“ als abhängige Variable erhoben (Eigenentwicklung). Nach der zufälligen Zuweisung zu den Treatmentgruppen wurde den Lernenden ein alltäglicher, besonderer oder innerfachlicher Kontext zugewiesen oder zur Auswahl bereitgestellt. Die Lernenden haben mithilfe ihres Kontextes an drei sequenzierten Lernaufgaben zum Inhalt „saure und alkalische Lösungen“ gearbeitet. Nach jeder Lernaufgabe wurden das situationale Interesse (Engeln, 2004) sowie die kognitive Belastung (Schwamborn et al., 2011) als abhängige Variablen gemessen. Zuletzt wurde das Fachwissen im Inhaltsfeld „saure und alkalische Lösungen“ erneut gemessen.

Ergebnisse

Die Modellierung durch ein Random-Intercept-Modell für das situationale Interesse verweist auf einen signifikanten Haupteffekt des Messzeitpunktes und des Treatments. Zu allen Messzeitpunkten zeigen sich signifikante Unterschiede zwischen der choose & match und der no choose & no match (b = 0.42, p < .001, r = .31) sowie der no choose & match und der no choose & no match Treatmentgruppe (b = 0.44, p < .001, r = .32). Zu keinem Messzeitpunkt besteht ein Unterschied zwischen der choose & match und der no choose & match Bedingung.

Ferner zeigen alle Lernenden am Ende der Intervention ein höheres Fachwissen als vorher. Die kognitive Belastung steigt signifikant von der ersten zur zweiten Aufgabe, bleibt dann aber konstant.

Bedeutung der Ergebnisse

Die Lernenden zeigen ein höheres situationales Interesse, wenn sie mit einem Kontext lernen, der zu ihren Personenmerkmalen passt. Hierbei ist es irrelevant, ob die Lernenden den Kontext selbst auswählen oder zugewiesen bekommen (z. B. Patall et al., 2008). Die Ergebnisse verdeutlichen damit, wie wichtig die passende Auswahl von Kontextaufgaben für den Chemieunterricht ist, um das situationale Interesse der Lernenden zu fördern.

Differentielle Wirksamkeit von Schwierigkeiten beim Experimentieren auf den individuellen Lernerfolg im Fach Chemie

Dennis Kirstein, Maik Walpuski
Universität Duisburg-Essen, Deutschland

Dem Experimentieren kommt im Chemieunterricht eine Schlüsselrolle für kompetenzorientiertes Lernen zu (Wirth et al., 2008). Damit liegt ein bedeutungsvoller Zugang für individualisierte Lernprozesse vor allem in einer bedarfsorientierten und adaptiven Gestaltung lernendenzentrierter Experimentierphasen (u.a. Hardy et al., 2011). Im Fach Chemie haben sich vor allem Experimentierformate bewährt, bei denen die Lernenden kooperativ an einer offenen Experimentieraufgabe arbeiten (Schulz, 2010; Rumann, 2005). Die bisherige Forschung zur Individualisierung im Fach Chemie liefert eher uneinheitliche Ergebnisse (vgl. Hauerstein, 2019; Anus, 2016), weist aber auf Lernvorteile durch diagnosegestützte Lernangebote hin (Michna, 2022). Für einzelne Inhaltsbereiche ist bereits bekannt, dass beim eigenständigen Experimentieren Lernschwierigkeiten und Fehler zu erwarten sind (u.a. Baur, 2018; Kechel, 2016) und individuelle Lernvoraussetzungen in einem engen Zusammenhang zum Lernerfolg stehen (Stender et al., 2018). Inwieweit sich beide Merkmale zur Individualisierung eignen und welche differentiellen Zusammenhänge bestehen, ist bislang noch nicht näher geklärt. In der vorgestellten Studie wird daher in einem Mixed-Methods-Ansatz untersucht, inwieweit unterschiedliche Schwierigkeiten spezifisch in individuellen Lernprozessen auf den Lernerfolg wirken. Dazu wurden die individuellen Lernvoraussetzungen und Bearbeitungsprozesse von 101 Schülerinnen und Schülern des zweiten Lernjahres Chemie in der Sekundarstufe I unterschiedlicher Schulformen beim Lernen mit kooperativen Experimentieraufgaben zu den lehrplanrelevanten Themen Batterien, Ozeanversauerung und Wasseranalyse mit Hilfe IRT-gestützter Methoden (Boone, Staver Yale, 2014) und strukturierender qualitativer Inhaltsanalyse (Mayring, 2015) untersucht. Das konzeptbezogene Fachwissen (Celik & Walpuski, 2020; eigene Entwicklung), das prozessbezogene Wissen zum Experimentieren (Mannel, 2009; Koenen, 2014) sowie die kognitiven Grundfähigkeiten (Heller & Perleth, 2000) als individuelle Lernvoraussetzungen wurden mit Leistungstests erfasst. Die Tests weisen für das konzept- ($Rel_{WLE} \geq 0{,}681$) und prozessbezogene ($Rel_{WLE} = 0{,}678$) Wissen sowie die kognitiven Grundfähigkeiten ($Rel_{WLE} = 0{,}858$) hinreichend hohe Reliabilitäten auf.

Die in den Bearbeitungsprozessen auftretenden Schwierigkeiten konnten mit Hilfe eines bereits erprobten Messverfahrens (Kirstein, Habig & Walpuski, 2020) hinreichend genau erfasst werden ($\kappa = 0{,}780$). Hierbei lassen sich themenunabhängig zwölf Schwierigkeiten beobachten: Schwierigkeiten in Bezug auf die Zielsetzung der Aufgabenstellung, Schwierigkeiten beim experimentellen Arbeiten, unsystematische Untersuchungen, Untersuchung nicht zielführender Zusammenhänge, Schwierigkeiten beim Umgang mit dem Experimentiermaterial, ungünstige Versuchsbedingungen,

inhaltliche und verständnisbezogene Fehler, fehlerhafter Umgang mit dem Experimentiermaterial, Fehler beim Schlussfolgern, Fehler bei der Untersuchung aufgabenbezogener Zusammenhänge, fehlerhafte Lösung der Aufgabe sowie unvollständige Lösung zur Aufgabe.

Zur Erfassung des Lernerfolgs wurde der Leistungstest zum konzeptbezogenen Fachwissen vor und nach der Bearbeitung der Experimentieraufgaben eingesetzt. Insgesamt lässt sich eine signifikante positive Entwicklung im konzeptbezogenen Fachwissen durch die Bearbeitung der Experimentieraufgaben beobachten ($t(100)=5{,}763$, $p<0{,}001$, $d=0{,}996$). Als wesentliche Bedingungsfaktoren für den individuellen Lernerfolg lassen sich mit Hilfe linearer Regressionsanalysen global für alle Schülerinnen und Schüler insbesondere das prozessbezogene Wissen zum Experimentieren ($F(2,99)=15{,}610$, $p<0{,}001$, $R^2=0{,}305$) sowie Schwierigkeiten in Bezug auf die Aufgabenstellung, Fehler beim Schlussfolgern sowie eine unvollständige Lösung der Aufgabe ($F(11,90)=4{,}332$, $p<0{,}001$, $R^2=0{,}286$) identifizieren.

Die Schülerinnen und Schüler können ausgehend von der Performanz im Leistungstests zum konzeptbezogenen Fachwissen grundsätzlich drei Leistungsgruppen zugeordnet werden, die sich hinsichtlich Lernerfolg und Lernvoraussetzungen unterscheiden: Schwache Lernende, durchschnittlich starke Lernende und überdurchschnittlich starke Lernende. Innerhalb der Leistungsgruppen treten die oben beschriebenen Schwierigkeiten in unterschiedlicher Intensität auf. Wird die oben beschriebene Analysen für die unterschiedlichen Leistungsgruppen wiederholt, deuten die Ergebnisse auf eine differentielle Wirksamkeit unterschiedlicher Schwierigkeiten hin. Während durchschnittlich starke Lernende in ihrem Lernerfolg durch unsystematische Untersuchungen ($F(1,20)=8{,}178$, $p=0{,}010$, $R^2=0{,}255$) beeinträchtigt werden, stellen für überdurchschnittlich starke Lernende vor allem eine unvollständige Lösung der Aufgaben sowie ein fehlerhafter Umgang mit dem Experimentiermaterial ($F(2,52)=25{,}298$, $p<0{,}001$, $R^2=0{,}474$) Risikofaktoren dar. Für schwache Lernender deutet sich ein negativer Einfluss von Schwierigkeiten in Bezug auf die Aufgabenstellung sowie durch Fehler bei der Untersuchung aufgabenbezogener Zusammenhänge an ($F(2,21)=2{,}674$, $p<0{,}092$, $R^2=0{,}127$). Auf der Tagung werden die Ergebnisse detailliert berichtet und hinsichtlich einer Eignung als Grundlage für eine Individualisierung von Lernprozessen beim Experimentieren diskutiert.

Risikokompetenz stärken – Der Einfluss von Informationsformat und Visualisierung bei der Erstellung von Diagrammen für zwei binäre Merkmale

Karin Binder[1], **Geier Corbinian**[2]
[1]Ludwigs-Maximilians-Universität München; [2]Adalbert-Stifter-Gymnasium Passau

In den Medien werden regelmäßig statistische FakeNews publiziert, die durch Fehlinterpretationen oder Verwechslungen bedingter Wahrscheinlichkeiten entstehen: Die Wahrscheinlichkeit einen SUV zu fahren, wenn man die Grünen wählt, wurde mit der Wahrscheinlichkeit grün zu wählen, wenn man einen SUV fährt, verwechselt (Unstatistik des Monats Mai 2021); der Erfolg der Tests der Gesichtserkennung am Bahnhof Südkreuz wurde drastisch überschätzt (Unstatistik des Monats Oktober 2018) und positive HIV-Tests hatten in der Vergangenheit sogar zu Suiziden geführt (Stine, 1996), obwohl manche dieser HIV-Tests wahrscheinlich nur fälschlicherweise positiv waren. All diese skizzierten Fehlurteile stellen sogenannte Bayesianische Situationen dar, in denen zwei Merkmale mit

je zwei Ausprägungen betrachtet werden (z.B. Gesichtserkennung schlägt Alarm vs. schlägt nicht Alarm; Person ist ein Gefährder vs. ist kein Gefährder).

Obwohl Bayesianisches Denken für viele Professionen von hoher Relevanz ist, belegt eine Vielzahl von Studien aus Kognitionspsychologie und Mathematikdidaktik die Schwierigkeiten in solchen Entscheidungsfindungsprozessen (Kahneman et al., 1982), aber auch mögliche Strategien, um diese Schwierigkeiten zu überwinden. Erfolgreiche Strategien sind hierbei: Die Kommunikation der statistischen Informationen mit natürlichen Häufigkeiten (z.B. 100 von 100.000 Personen) statt Wahrscheinlichkeiten (z.B. 0,1 %; siehe Gigerenzer & Hoffrage, 1995; McDowell & Jacobs, 2017) und die Visualisierung der statistischen Informationen (Binder, Krauss & Bruckmaier, 2015; Böcherer-Linder & Eichler, 2019).

Die bisherigen Studien zum Effekt der Visualisierungen bei Bayesianischen Denken gehen hierbei von bereits vollständig ausgefüllten Visualisierungen aus und untersuchen beispielsweise den Effekt beim *Ablesen* entsprechender Informationen (z.B. McDowell & Jacobs, 2017). In Lehr-Lern-Prozessen ist jedoch auch entscheidend, welchen Einfluss das Informationsformat und die Visualisierung beim *Anfertigen* der entsprechenden Visualisierung haben. Die vorliegende Studie fokussiert hierbei das Eintragen von statistischen Informationen in Vierfeldertafeln, Doppelbäume und Netzdiagramme (vgl. Kunzelmann et al., 2022). Dabei sollten 120 Schüler:innen (Sekundarstufe II) und 52 Studierende (Fachrichtung Biologie) in einem Papier-und-Bleistift-Test bedingte Wahrscheinlichkeiten (jeweils entweder im Format von Wahrscheinlichkeiten oder natürlicher Häufigkeiten) in leere vorgegebene Diagrammstrukturen (Vierfeldertafeln, Doppelbäume oder Netze) eintragen (die lediglich die Beschriftung der Ereignisse bereits beinhalteten). Zentrale Fragestellung der Studie ist der Einfluss von Informationsformat und Visualisierung auf die Performanz und auf typische Fehler beim Eintragen dieser Zahlen.

Die Ergebnisse zeigen bereits deskriptiv einen Effekt des Formats: Über die Visualisierungen hinweg werden natürliche Häufigkeiten (68%) häufiger korrekt eingetragen als Wahrscheinlichkeiten (54% - Vierfeldertafeln wurden hierbei ausgeschlossen, da hier bedingte Wahrscheinlichkeiten nicht direkt eingetragen werden können). Im Detail zeigen sich für die Häufigkeitsvarianten die folgenden Lösungsraten: 68% Vierfeldertafel, 61% Doppelbaum und 75% Häufigkeitsnetz. In den Wahrscheinlichkeitsvisualisierungen haben 61% der Versuchspersonen die Wahrscheinlichkeiten in den Doppelbaum richtig eingetragen und 47% in das Häufigkeitsnetz. Nur 12% der Versuchspersonen haben erkannt, dass ein direktes Eintragen einer bedingten Wahrscheinlichkeit in eine schulübliche Vierfeldertafel nicht möglich ist.

Inferenzstatistisch wird ein verallgemeinertes lineares gemischtes Modell mit logistischer Linkfunktion zur Vorhersage der Lösewahrscheinlichkeit und dem Häufigkeitsnetz als Referenzkategorie betrachtet. Dabei wurden das Informationsformat sowie die Visualisierungen Vierfeldertafel und Häufigkeitsnetz als Dummy-Variablen implementiert. Es zeigt sich ein (nicht-standardisierter) signifikanter Formateffekt ($b_1 = -1{,}25$; $p < 0{,}01$). Wahrscheinlichkeiten wurden also signifikant weniger häufig korrekt in das Netz eingetragen als Häufigkeiten. Die Visualisierungen stellen keine signifikanten Prädiktoren dar, was bedeutet, dass in alle drei Häufigkeitsvarianten ähnlich häufig die

Zahlen korrekt eingetragen wurden. Allerdings zeigen sich signifikante Interaktionseffekte von Vierfeldertafel×Format (b_2 = -1,66; p = 0,01) und Doppelbaum×Format (b_2 = 1,22; p = 0,04). Das bedeutet, dass der Formateffekt bei Vierfeldertafeln sogar noch stärker ausgeprägt ist, bei den Doppelbäumen hingegen verschwindet.

Der vielfach belegte Häufigkeitseffekt im Umgang mit bedingten Wahrscheinlichkeiten, zeigt sich also nicht nur beim Ablesen aus Diagrammen, sondern sogar beim Eintragen in Diagramme, wobei Doppelbäume hier eine Ausnahme bilden. Überdies liefert die Studie erste Befunde welche Visualisierungen welche typischen Fehler beim Eintragen von bedingten Wahrscheinlichkeiten provozieren oder aber davor bewahren.

Symposien

H-11: Qualität und Wirkung von Mathematikunterricht – Innovative Forschungsansätze auf der Basis der TALIS-Videostudie

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* S07 S00 D07

Qualität und Wirkung von Mathematikunterricht – Innovative Forschungsansätze auf der Basis der TALIS-Videostudie

Chair(s): **Eckhard Klieme** (DIPF | Leibniz Institut für Bildungsforschung und Bildungsinformation, Deutschland), **Anna-Katharina Praetorius** (Universität Zürich)
*Diskutant*in(nen):* **Anke Lindmeier** (Universität Jena)

Die deutschsprachige Unterrichtsforschung hat sich im Rahmen der internationalen TIMSS-Videostudien 1995/99 differenzierte Instrumentarien zur Videoanalyse erschlossen und u.a. das Konzept der drei Basisdimensionen der Unterrichtsqualität eingeführt (vgl. Praetorius et al., 2020a). In der Folge wurden videobasierte Wirkungsstudien implementiert, die jeweils eine Unterrichtseinheit in den Blick nahmen (z.B. Seidel et al. 2006; Klieme et al. 2006). Dieses Design konnte 2018 wiederum international umgesetzt werden: Die TALIS-Videostudie erfasste Unterrichtseinheiten zum Thema quadratische Gleichungen in acht Staaten (Opfer et al., 2020). Alle quantitativen Daten (Vor- und Nachtests, Befragungen, Ratings von Videos und Unterrichtsmaterialien) sind öffentlich verfügbar. Für Deutschland mit n=50 Klassen stehen der sekundäranalytischen Forschung auch Unterrichtsvideos und –materialien sowie nationale Ergänzungen (z.B. Follow Up sowie Zusatzvideos außerhalb der Unterrichtseinheit) zur Verfügung (Grünkorn et al., 2020).

Das Symposium demonstriert den Nutzen dieses Datenkorpus in vier Beiträgen, die im Rahmen von Qualifikationsarbeiten aus Erziehungswissenschaft und Pädagogischer Psychologie Daten der deutschen Stichprobe verwenden. Die Diskussion bringt zudem eine mathematikdidaktische Perspektive ein.

Alle Beiträge setzen an dem Modell der Basisdimensionen von Unterrichtsqualität an, d.h. an dem Strukturmodell von Klassenführung, Konstruktiver Unterstützung und Kognitiver Aktivierung sowie an dem Wirkungsmodell, wonach die Qualitätsdimensionen vermittelt über time-on-task, Verarbeitungstiefe und die Erfüllung von Grundbedürfnissen im Sinne der Selbstbestimmungstheorie auf Fachleistung und Fachinteresse wirken (vgl. Praetorius et al., 2020b, zum aktuellen Stand). Die Beiträge geben Impulse zur Überprüfung des Modells sowie für unterrichtstheoretische und forschungsmethodische Fragestellungen:

- Der erste Beitrag prüft die im Wirkungsmodell formulierten Effekte einschließlich der Mediationshypothesen anhand von Mehr-Ebenen-Analysen. Grundlage sind dabei die deutschen Schülerdaten der TALIS-Videostudie, insbesondere die Wahrnehmung der Unterrichtsqualität, selbstberichtete Nutzungsaspekte sowie Leistung und Fachinteresse.
- Der zweite Beitrag operationalisiert die Basisdimensionen mit Hilfe der international standardisierten Video-Ratings. Diese werden einerseits auf die nationalen Zusatzvideos aus dem Jahr 2018 angewandt, andererseits auf Videos aus TIMSS 1995. Die Unterrichtsforschung wird somit um eine (zeit)historische Perspektive erweitert, die den Wandel von Unterrichtskulturen nachzeichnet. Zudem wird das Strukturmodell der Basisdimensionen faktorenanalytisch untersucht und die Messinvarianz zwischen 1995 und 2018 geprüft.
- Der dritte Beitrag konfrontiert die standardisierten Ratings trainierter Beobachter*innen mit neuen Einschätzungen, die untrainierte Beobachter*innen anhand von kurzen Videoclips („Thin Slices“) abgegeben haben. Ziel ist der Nachweis, dass die Thin-Slices-Technik zu durchaus validen Einschätzungen der Unterrichtsqualität führt, die auch Schüleroutcomes erklären.
- Der vierte Beitrag fokussiert die komplexeste der drei Dimensionen, die Kognitive Aktivierung. Er bringt die standardisierten Ratings hierzu mit qualitativ-rekonstruktiven Interpretationen zusammen. Ziel ist eine Triangulation qualitativer und quantitativer Ansätze.

Beiträge des Symposiums

Der Zusammenhang zwischen Unterrichtsqualität und Schüleroutcomes: Die vermittelnde Rolle der Nutzung von Lerngelegenheiten

Aysenur Alp Christ[1], Vanda Capon-Sieber[1], Carmen Köhler[2], Anna-Katharina Praetorius[1]
[1]Universität Zürich, [2]DIPF | Leibniz Institut für Bildungsforschung und Bildungsinformation, Deutschland

Theoretischer Hintergrund

Das Modell der drei Basisdimensionen der Unterrichtsqualität (TBD) (Klieme et al., 2006; Praetorius et al., 2018) gilt als zentrales Modell in der Unterrichtsqualitätsforschung. Mit dem Ziel, einen Überblick über Aspekte der Unterrichtsqualität zu geben, unterschieden Klieme et al. (2001) zwischen kognitiver Aktivierung (z.B. Stellen nicht-routinemässiger Probleme), Klassenführung (z.B. Vermeidung von Störungen) und konstruktiver Unterstützung (z.B. Autonomieunterstützung) und lieferten erste empirische Belege für deren Zusammenhang mit Schülerleistung. Konstruktivistischen Annahmen folgend wurden nachträglich Mediatoren zwischen Unterrichtsqualität und Schüleroutcomes hinzugefügt, um zu erklären, wie Unterrichtsqualität die Nutzung der Lerngelegenheiten durch Schüler*innen beeinflusst (z. B. Klieme et al., 2006). Die Annahmen darüber, welche Mediatoren zu erwarten sind und welche nicht, wurden durch den Bezug auf einige wenige ausgewählte theoretische Ansätze, unter anderem die Selbstbestimmungstheorie (Ryan & Deci, 2000), geleitet. So ist eine Annahme im Modell, dass konstruktive Unterstützung über die Bedürfnisbefriedigung auf die Motivation wirkt. Dieses Vorgehen könnte jedoch dazu geführt haben, dass andere, theoretisch ebenfalls plausible Vermittlungsmechanismen unbeachtet blieben. Wird beispielsweise die im Feld ebenfalls prominente Erwartungs-Wert-Theorie (Wigfield et al., 1992) hinzugenommen, erscheinen weitere Mediationspfade plausibel. Laut Erwartungs-Wert-Theorie beeinflussen Aspekte der kognitiven Aktivierung, wie optimale Herausforderung, die Erfolgserwartungen der Schüler*innen sowie ihr Kompetenzerleben. Somit scheint ein Zusammenhang zwischen kognitiver Aktivierung und Bedürfnisbefriedigung plausibel, der im ursprünglichen Modell jedoch nicht berücksichtigt wurde. Eine umfassende Literaturrecherche zeigt auf, dass unter Einbezug weiterer Motivations- und Kognitionstheorien inhaltliche Argumente für Zusammenhänge zwischen jeder Unterrichtsqualitätsdimension und den drei im Modell postulierten Mediatoren (Verarbeitungstiefe, time-on-task und Bedürfnisbefriedigung) bestehen. Die Problematik wird zudem dadurch verstärkt, dass, obwohl die Mediationsannahmen für das Modell zentral sind, diese nur selten empirisch überprüft wurden. Das Ziel der vorliegenden Studie ist deshalb einerseits die empirische Überprüfung des TBD-Modells und andererseits die Überprüfung weiterer theoretisch schlüssiger Vermittlungspfade.

Methode:

Es wurden Schülerwahrnehmungen von n=958 Gymnasiast*innen der TALIS-Videostudie 2018 in Deutschland verwendet. Diese wurden an zwei Messzeitpunkten in 41 Klassen erhoben (Malter=14,82, SD=0,62; 50,5% weiblich). Die Unterrichtsqualitätsdimensionen kognitive Aktivierung (sieben Items), Klassenführung (acht Items) und konstruktive Unterstützung (elf Items) fungierten jeweils als unabhängige Variable. Die Mediatoren Verarbeitungstiefe, time-on-task und Bedürfnisbefriedigung wurden mit jeweils drei Items erfasst. Die abhängige Variable Interesse wurde ebenfalls mit drei Items erfasst. Die abhängige Variable Mathematikleistung wurde mit 30 Aufgaben

erhoben. Für die Durchführung von Mehrebenenanalysen im Längsschnitt wurde Lavaan (Rosseel, 2012) in R (R Core Team, 2022) verwendet. Zunächst wurde der Zusammenhang zwischen jeder Unterrichtsqualitätsdimension und Leistung und Interesse modelliert. Anschliessend wurde mittels drei partiellen Mediationsmodellen der indirekte Effekt jeder Unterrichtsqualitätsdimension auf Leistung und Interesse untersucht, vermittelt durch die drei Mediatoren (Verarbeitungstiefe, time-on-task und Bedürfnisbefriedigung) unter Kontrolle von Leistung und Interesse.

Ergebnisse und Diskussion:

Weder auf Klassen- noch auf Individualebene konnten Effekte der Unterrichtsqualitätsmerkmale auf die Schüleroutcomes festgestellt werden. Es fanden sich hohe Zusammenhänge zwischen den Unterrichtsqualitätsmerkmalen und den Mediatoren, die partiellen Mediationsmodelle auf Klassenebene konnten die Mediationsannahmen jedoch nicht bestätigen. Auf Schülerebene zeigten sich zwei mit dem TBD-Modell übereinstimmende Mediationen: Verarbeitungstiefe mediierte den Zusammenhang zwischen kognitiver Aktivierung und Leistung. Weiter mediierte Bedürfnisbefriedigung den Zusammenhang zwischen konstruktiver Unterstützung und Interesse. Im Gegensatz zum TBD-Modell und im Einklang mit unseren theoretischen Annahmen, wurden weitere Zusammenhänge gefunden: Time-on-task mediierte den Zusammenhang zwischen Klassenführung und Interesse sowie zwischen konstruktiver Unterstützung und Interesse. Verarbeitungstiefe mediierte den Zusammenhang zwischen konstruktiver Unterstützung und Leistung. Unsere Ergebnisse zeigen, dass das TBD-Modell von einer Erweiterung der Annahmen zu den Mediatoren auf der Grundlage relevanter theoretischer Ansätze profitieren könnte. Darüber hinaus sollte die TBD-Forschung die Möglichkeit zusätzlicher Mediatoren (z.B. Erfolgserwartung) in Betracht ziehen. Die Ergebnisse auf Individual- und Klassenebene sind unterschiedlich. Künftige Studien sollten sich eingehend der Frage widmen, auf welcher Ebene Mediationseffekte zu erwarten sind.

Die drei Basisdimensionen von Unterrichtsqualität: Was hat sich in den letzten zwanzig Jahren verändert?

Leonie Gossner[1], Wida Wemmer-Rogh[1], Patrick Schreyer[2], Urs Grob[1], Eckhard Klieme[3], Anna-Katharina Praetorius[1]
[1]Universität Zürich, [2]Universität Kassel, [3]DIPF | Leibniz Institut für Bildungsforschung und Bildungsinformation, Deutschland

Theoretischer Hintergrund und Fragestellung

Historische Veränderungen von Schule und Unterricht im Verlauf des letzten Jahrhunderts wurden bislang primär bildungsgeschichtlich und qualitativ-empirisch untersucht (Feldhoff et al., 2020). Tyack und Tobin (1994) gehen etwa von einer ‘Grammar of Schooling’ aus, die über Jahrzehnte stabil und unveränderlich bleibende institutionelle Rahmenbedingungen von Schule beschreibt und wie ein Filter gegenüber Veränderungsversuchen wirkt. Inwiefern auch die konkrete Unterrichtspraxis in ihren Grundzügen über längere Zeiträume invariabel bleibt, ist unklar. Hinweise auf langfristige Veränderungen in der Unterrichtsinteraktion zwischen Lehrpersonen und den Schüler*innen geben Untersuchungen, die von Transformationsprozessen hin zu mehr Schüler*innenorientierung und -mitbestimmung berichten (z. B. Caruso, 2003; Depaepe et al., 2000; Helsper, 2000). Keine dieser Studien konnte jedoch direkte Beobachtungen von Unterrichtssituationen nutzen, sie weisen somit Einschränkungen in ihrer Erklärungskraft hinsichtlich der konkreten Unterrichtsprozesse auf (Caruso & Reh, 2021).

Zur Untersuchung ebendieser Prozesse hat sich in den vergangenen zwei Dekaden besonders im deutschsprachigen Raum das Modell der drei Basisdimensionen von Unterrichtsqualität etabliert (Praetorius et al., 2018). Die Basisdimensionen – Klassenführung, konstruktive Unterstützung und kognitive Aktivierung – widerspiegeln «die unverzichtbaren Grundbedingungen schulisch organisierten Lernens» (Klieme et al., 2001, S. 52). Seit der Erschließung der drei Basisdimensionen im Rahmen der deutschen TIMSS-Videostudie 1995 wurden etliche Versuche unternommen, die Qualität von Unterricht weiterzuentwickeln (z. B. SINUS-Modellversuchsprogramm; Prenzel, 2019). Auch qua Lehrpersonenaus- und weiterbildung dürften Erkenntnisse der empirischen Unterrichtsqualitätsforschung in die Praxis eingeflossen sein. So ließe sich vermuten, dass die Unterrichtsqualität im Vergleich zu den 90er Jahren heute qualitativ höher ausgeprägt ist. Ob sich Hinweise darauf zeigen, wird in vorliegender Studie untersucht.

Methode

Mit der TALIS-Videostudie Deutschland 2018 (Grünkorn et al., 2020), die Parallelen zur deutschen TIMSS-Videostudie 1995 aufweist, bietet sich eine seltene Gelegenheit, dieser Frage nachzugehen. Die erweiterte TALIS-Videostudie Deutschland umfasst Videografien des Mathematikunterrichts zu unterschiedlichen Mathematikthemen der 8.-10. Klassenstufen (n = 44). Zum Zweck des historischen Vergleichs wurde aus der TIMSS-Videostudie 1995 der 8. Klassen eine annähernd vergleichbare Teilstichprobe (n = 44) bezüglich der Schularten, Bundesländer sowie behandelter Unterrichtsinhalte gebildet. Alle Unterrichtsvideos (N = 88) wurden von geschulten Rater*innen anhand des TALIS-Beobachtungssystems eingeschätzt (Bell, 2020a). Die hoch-inferente Beurteilung der Unterrichtsaufzeichnungen erfolgte mit einer vierstufigen, verhaltensverankerten Ratingskala (Bell, 2020a) und je drei bis sechs Indikatoren pro Basisdimension (Klieme & Schreyer, 2020). Die durchschnittliche exakte Rater*innenübereinstimmung liegt in der TIMSS-Teilstichprobe bei 57 % und der TALIS-Teilstichprobe bei 56 %, was nach Bell (2020b) einer hinreichenden Übereinstimmung für hoch-inferente Beobachtungsmerkmale entspricht. Anschließend erfolgte ein für die Basisdimensionen getrenntes Vorgehen zur Prüfung von Messinvarianz zwischen den Teilstichproben.

Ergebnisse und ihre Bedeutung

Bei der Dimension kognitive Aktivierung ließ sich unter Ausschluss eines Indikators auf der Grundlage einer zweifaktoriellen Lösung skalare Messinvarianz etablieren, was Vergleichsmöglichkeiten in Form von latenten Mittelwertunterschieden ermöglichte. Demgegenüber gestaltete sich die Beziehungsstruktur der Indikatoren von Klassenführung über die Teilstichproben weit heterogener, sodass keine Messinvarianz aufgezeigt werden konnte. Auch bei der konstruktiven Unterstützung liegt keine geteilte faktorielle Grundstruktur in den Teilstichproben und somit keine Messinvarianz vor.

Diese ersten Ergebnisse und insbesondere die fehlende Messinvarianz bei Klassenführung und konstruktiver Unterstützung liefern Hinweise darauf, wie sich Unterrichtsprozesse zwischen der deutschen TIMSS-Videostudie und der TALIS-Videostudie Deutschland gewandelt haben. Dass bei kognitiver Aktivierung als einziger Dimension Messinvarianz vorliegt, erscheint zunächst kontraintuitiv, da bisherige Studien auf deren Instabilität und Kontextabhängigkeit hingewiesen haben (z. B.

Praetorius et al., 2014). Mögliche Erklärungsansätze für die divergierenden Ergebnisse bezüglich der Messinvarianz liegen einerseits in statistischen Artefakten aufgrund geringer Varianzen, andererseits auf inhaltlicher Ebene in differenziellen Entwicklungen der Unterrichtsgestaltung sowie einer veränderten Beziehungsdynamik zwischen Lehrpersonen und Schüler*innen (vgl. Fend & Berger, 2016). Diesen Anschlusshypothesen soll in vertiefenden Analysen nachgegangen werden. Die Datenlage mahnt jedoch aufgrund der Nicht-Repräsentativität der Stichproben, der geringen Fallzahl sowie der mitunter relativ niedrigen Rater*innenübereinstimmung zu Vorsicht bei Generalisierungen.

Sind 30 Sekunden genug? Analysen zur Reliabilität und Validität von Thin-Slices Ratings bei der TALIS-Videostudie

Konstantin Vinokic[1], Lukas Begrich[2], Mareike Kunter[2], Susanne Kuger[1]
[1]Deutsches Jugendinstitut, [2]DIPF | Leibniz Institut für Bildungsforschung und Bildungsinformation, Deutschland

Theoretischer Hintergrund

Insbesondere die Qualität der lernbezogenen Interaktionen im Klassenraum ist dafür entscheidend, wie viel Schüler:innen im Unterricht lernen und wie viel Freude und Interesse sie am Lernen entwickeln. Die Qualität von Unterricht lässt sich anhand der drei Basisdimensionen kognitive Aktivierung, strukturierte Klassenführung und konstruktive Unterstützung beschreiben (Klieme et al., 2009; Kunter & Voss, 2011). Zur Erfassung dieser Qualitätsmerkmale wird häufig auf Videoratings externer Beobachter:innen zurückgegriffen, die zumeist auf Grundlage einzelner Unterrichtsstunden abgeben werden. Dieser Ansatz ist sowohl kosten- als auch zeitintensiv, da die Beobachter:innen aufwändig geschult werden müssen (Helmke, 2014; Janik & Seidel, 2014).

Thin-Slices Ratings sind Einschätzungen persönlicher Eigenschaften oder Verhaltensmerkmalen von Zielpersonen, die lediglich auf ersten Eindrücken ungeschulter Beobachter:innen beruhen. Ein Thin-Slice ist ein Exzerpt aus dem dynamischen Verhalten von Personen, von höchstens 5 Minuten Länge (Ambady, et al., 2000). Zahlreiche Studien belegen sowohl die Reliabilität solcher Einschätzungen im Sinne hoher Übereinstimmungen zwischen Beobachter:innen als auch ihre Validität im Sinne bedeutsamer Zusammenhänge mit Kriterien wie etwa Selbstberichten der eingeschätzten Zielpersonen (z.B. Ambady et al., 2000).

Fragestellung

Die Genauigkeit der Thin-Slices Ratings wirft die Frage auf, ob diese auch dazu genutzt werden können, die Unterrichtsqualität reliabel und valide zu erfassen. Die Ergebnisse der Studien von Begrich et al. (2017, 2020, 2021) konnten zeigen, dass Thin-Slices Einschätzungen in Einzelfällen auch für die Erfassung von Unterrichtsqualität geeignet sind. Urteile auf Grundlage 30-sekündiger Videoausschnitte stimmen 1. hoch miteinander überein (Reliabilität), 2. hängen teilweise signifikant mit Ratings geschulter Beobachter:innen von vollständigen Unterrichtsvideos zusammen (konvergente Validität), 3. erlauben eine differenzierte Erfassung der drei Basisdimensionen (Konstruktvalidität) und 4. sind prädiktiv im Hinblick auf den Lernerfolg von Schüler:innen (prädiktive Validität). Das Ziel der vorliegenden Arbeit ist die Prüfung der Generalisierbarkeit der Befunde von Begrich und Kolleg:innen (2017, 2020,2021).

Methode

In der vorliegenden Arbeit wurde eine Re-analyse der TALIS-Videostudie Deutschland (2020) durchgeführt. Als Grundlage dienten 30-sekündige, audiovisuelle Slices, die randomisiert aus jedem Drittel jeweils einer Unterrichtsstunde gesampelt wurden. Diese Slices wurden von 11 ungeschulten Rater:innen anhand von 25 Items zu den drei Basisdimensionen (Klieme et al., 2009; Kunter & Voss, 2011) eingeschätzt. Die resultierenden Daten wurden systematisch mit den Videoanalysen der TALIS-Videostudie verglichen sowie genutzt, um die Prädiktionsmodelle der Studie bzgl. verschiedener Schüleroutcomes zu replizieren.

Ergebnisse und ihre Bedeutung

Die Beurteilerübereinstimmungen zu den 25 Items waren heterogen. Die Reliabilität der Ratings (ICC2) für die Items zur Klassenführung reichten von .60 bis .74 , derer zur Konstruktiven Unterstützung reichten von .25 bis .80 und derjenigen für die Kognitive Aktivierung reichten von .18 bis .63. Anhand der ICC wurden die Skalen optimiert, indem Items unter .50 entfernt wurden. Der relative G-Koeffizient über alle drei Basisdimensionen war mit .61 kleiner als der akzeptable Wert von ≥ .70 (Gabriel-Busse & Lipowsky, 2020). Der relative G-Koeffizient der Klassenführung war .85, der Konstruktiven Unterstützung war .75 und der Kognitiven Aktivierung war .55.

Zur Überprüfung der konvergenten Validität wurden Thin-Slices Ratings der Unterrichtsqualität mit den Ratings geschulter Rater:innen auf Grundlage vollständiger Unterrichtsvideos verglichen. Es zeigten sich systematisch konvergente Korrelationen. Die prädiktive Validität der Thin-Slices Ratings der Unterrichtsqualität hinsichtlich verschiedener Schüleroutcomes wurde anhand von Mehrebenenanalysen überprüft. Thin-Slices Ratings der Unterrichtsqualität konnten die Schülerleistungsentwicklung nicht vorhersagen. Jedoch konnten Thin-Slices Ratings der Kognitiven Aktivierung und teilweise der Konstruktiven Unterstützung signifikant sowohl die Entwicklung des Schülerinteresses als auch die der Selbstwirksamkeit vorhersagen. Trotz mäßiger Skalenreliabilität konnten sowohl Schüleroutcomes signifikant vorhergesagt werden als auch signifikante Übereinstimmungen mit den Ratings geschulter Rater:innen auf Grundlage vollständiger Videos erzielt werden. Die Befunde implizieren, dass Thin-Slices Ratings der Unterrichtsqualität valide Ergebnisse liefern können.

Unterrichtsqualität in der Triangulation von quantitativen und qualitativen Zugängen: Kognitive Aktivierung in Ratings und Fallrekonstruktionen

Patrick Schreyer[1], Eckhard Klieme[2]
[1]Universität Kassel, [2]DIPF | Leibniz Institut für Bildungsforschung und Bildungsinformation, Deutschland

Theoretischer Hintergrund und Fragestellung

Die Unterrichtsqualitätsforschung befasst sich primär mit der empirischen Untersuchung von Lehr- und Lernprozessen im Unterricht und untersucht die Wirksamkeit bestimmter Merkmale auf die Lernleistung von Schüler*innen (Praetorius et al., 2020; Rakoczy & Klieme, 2016). Um den Lernerfolg im Unterricht untersuchen zu können, kommen bei diesem als nomothetisch zu charakterisierenden Forschungsstrang überwiegend quantitative Methoden, wie bspw. Fragebogenerhebungen oder standardisierte Beobachtungen, zum Einsatz. Parallel dazu bestehen zahlreiche qualitativ-rekonstruktive Forschungsansätze, die sich mit der sozialen Verfasstheit sowie den Gelingensbedingungen von unterrichtlichen Interaktionen beschäftigen (Proske et al., 2021).

Wenngleich bei diesen Ansätzen weniger die Frage nach der ‚qualitativen' Gestaltung bzw. der Effektivität des unterrichtlichen Geschehens im Vordergrund der Analysen steht (Praetorius et al., 2021), so lassen sich zunehmend Arbeiten finden, die konkret nach den Lern- und Verstehensprozessen des Unterrichts fragen (Breidenstein & Tyagunova, 2020; Hackbarth et al., 2022). Obgleich sich beide Forschungsansätze für das Lernen von Schüler*innen im Unterricht interessieren, stehen die quantitative und die qualitative Unterrichtsforschung bis heute relativ unabhängig nebeneinander (Proske & Rabenstein, 2018). In der Erforschung von Unterrichtsqualität existieren bislang nur wenige Versuche einer theoretischen oder methodischen Verknüpfung beider Stränge (bspw. Klieme & Bos, 2000; Praetorius et al., 2021; Vieluf & Klieme, im Druck). Innerhalb dieses Beitrags soll der Versuch einer Triangulation zweier unterschiedlicher, methodischer Zugänge vorgenommen werden. Am Beispiel des Merkmals der kognitiven Aktivierung (Klieme, 2019) werden die Ergebnisse einer quantitativen und einer qualitativ-rekonstruktiven Einschätzung der Unterrichtsqualität ausgewählter und auf Video aufgezeichneter Unterrichtsstunden gegenübergestellt und miteinander verglichen. Das Ziel dieses Beitrags ist es, durch die wechselseitige Betrachtung des gleichen Datenmaterials eine Erweiterung der Erkenntnismöglichkeiten (Denzin, 2009) und die Generierung eines reichhaltigeren Bildes des untersuchten Gegenstands (Kelle, 2008), sprich ein differenzierteres Verständnis des Konstrukts ‚kognitive Aktivierung', zu generieren.

Methode

Ausgangspunkt des triangulierenden Vergleichs sind die Unterrichtsaufzeichnungen aus dem Mathematikunterricht von zwei Gymnasiallehrkräften mit ihren Klassen der Stufe 9 aus der TALIS-Videostudie Deutschland (Grünkorn et al., 2020). In beiden Unterrichtsstunden wird die gleiche Aufgabe gestellt und bearbeitet: Die Schüler*innen sind angehalten, den maximalen Flächeninhalt eines Kaninchengeheges bei einer vorgegebenen Zaunlänge zu berechnen. Ausgehend von dieser Modellierungsaufgabe im Themenbereich der quadratischen Gleichungen wird ein Vergleich zwischen den beiden Fällen mit unterschiedlichen methodischen Zugängen vorgenommen. Im Sinne einer ‚methodischen Triangulation' (Denzin, 2009) zweier unterschiedlicher Methoden soll der Vergleich primär anhand der durch die beiden Methoden erzielten Ergebnisse erfolgen (Kelle, 2008). Dabei wird im Beitrag auf zwei stark kontrastive Ansätze zurückgegriffen, die bislang noch nicht in dieser Form miteinander in Beziehung gesetzt wurden. Beim ersten Ansatz wurden die Unterrichtsvideos quantitativ mit dem in der TALIS-Video Study (OECD, 2020) entwickelten Kodiermanual beurteilt (Bell et al., 2020; Klieme & Schreyer, 2020). Parallel und unabhängig von diesen Auswertungen wurden die Videos mithilfe eines qualitativ-rekonstruktiven Verfahrens, der Dokumentarischen Methode (Asbrand & Martens, 2018; Bohnsack, 2021), mit dem Fokus auf kognitive Aktivierung, z.B. interaktive Bearbeitung von Aufgabenimpulsen, ausgewertet.

Ergebnisse und ihre Bedeutung

Durch die Verknüpfung der Ergebnisse der beiden methodischen Zugänge lassen sich sowohl konvergierende, komplementäre als auch divergente Befunde herausarbeiten (Kelle & Erzberger, 2003). Konvergierend zeigt sich, dass jeweils beide Ansätze, die quantitativen Einschätzungen und die qualitativen Rekonstruktionen, bei der gleichen Unterrichtsstunde eine höhere kognitiv

aktivierende Bearbeitung der Aufgabe konstatieren. Komplementär geben die Befunde zu erkennen, dass mithilfe der quantitativen Beobachtungsurteile verstärkt die Relationierung von vorab festgelegten Merkmalen zu anderen Klassen ermöglicht wird, während durch die dokumentarische Analyse stärker die Verschränkungen und die Zusammenhänge von identifizierten (Qualitäts)Merkmalen herausgearbeitet werden können. Gleichzeitig vermag die qualitativ-rekonstruktive Analyse die konkrete Bearbeitung des Eingangs gestellten Aufgabenimpulses zu beschreiben, während die standardisierte Beobachtung vermehrt eine aufgabenunabhängige Qualität des Unterrichts in den Blick nimmt. Besonders diese divergenten Befunde können Anlass für weitere Diskussionen und die Ausdifferenzierung der Konzeptionalisierung von Unterrichtsqualität bieten.

H-13: Fachspezifische Lehrkräftekompetenzen im Erklären: Überblick und Ergebnisse der drei Projekte FALKE-q, -e, -d

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* T03 R02 D39

Fachspezifische Lehrkräftekompetenzen im Erklären: Überblick und Ergebnisse der drei Projekte FALKE-q, -e, -d

Chair(s): **Anita Schilcher** (Universität Regensburg, Deutschland), **Stefan Krauss** (Universität Regensburg, Deutschland)
*Diskutant*in(nen):* **Markus Vogel** (PH Heidelberg), **Christoph Kulgemeyer** (Universität Bremen)

Studien wie etwa diejenige von Wragg und Wood (1984) konnten zeigen, dass die Fähigkeit, gut erklären zu können, von Schülerinnen und Schülern als die wichtigste Kompetenz der Lehrkräfte gesehen wird (vgl. auch Rathausky, 2010; Wilson & Mant, 2011). Im Kontext der Forschung zu professionellen Kompetenzen lässt sich das Konstrukt der Erklärkompetenz theoretisch als Teilfacette fachdidaktischen Wissens bzw. des pedagogical content knowledge (PCK) von Lehrkräften verorten (Shulman, 1987, Krauss et al., 2017). In der jüngeren Vergangenheit wird „Erklären“ als eine *core practice* (McDonald, M., Kazemi, E. & Kavanagh S. H., 2013) beschrieben, in der sich psychologische, pädagogische, fachliche und fachdidaktische Wissensbestände zu einer komplexen Kompetenz verbinden (Sevian & Gonsalves, 2008; Findeisen, 2017). Als performative Handlungskompetenz ist das Erklären Teil der professionellen Gesprächskompetenz von Lehrkräften und ein wichtiges Element für das Verstehen komplexer Zusammenhänge. Es stellt sich deshalb auch die Frage, ob der stetig anwachsende Markt an Erklärvideos (Rat für kulturelle Bildung, 2019) kompensatorisch wirken kann.

Trotz der Bedeutung des schulischen Erklärens ist die Forschungslage zum Erklären insgesamt defizitär. Das liegt unter anderem daran, dass nicht geklärt ist, ob es sich bei der Erklärkompetenz um eine primär fachliche oder generische Kompetenz handelt. Erklären gehört weder in den Erziehungswissenschaften noch in den Fachdidaktiken zum Standardrepertoire der Ausbildung (Schilcher et al., 2017). Deswegen hat sich das transdisziplinäre Forschungsprogramm FALKE dem Erklären gewidmet, dessen Einzelprojekte in vier Vorträgen präsentiert werden.

- In FALKE-q (quality) steht die empirische Untersuchung von Kriterien guten Erklärens im Mittelpunkt (Schilcher et al., i.V.). Dabei werden Gemeinsamkeiten und Unterschiede zwischen einzelnen Schulfächern dargestellt.
- Darauf aufbauend untersucht FALKE-e (expertise) mit einer Interventionsstudie, inwiefern Erklärkompetenz bei Studierenden gezielt trainiert werden kann. Dabei wird in einem Beitrag dargestellt, inwiefern die Beurteilung von Erklärungen durch Schüler:innen mit deren Leistung in einem Zusammenhang steht. Ein zweiter Beitrag zeigt, wie sich Selbstwirksamkeitserwartungen und Unsicherheitstoleranz durch die mehrfache Reflexion eigener Erklärungen verändern.
- In FALKE-d (digital) werden Bedingungen der Wirksamkeit von Erklärvideos im Flipped-Classroom untersucht. Vergleichen werden drei Treatments: eine Gruppe mit Strategietraining für das Rezipieren von Erklärvideos im flipped-classroom, eine Gruppe ohne Strategietraining und regulärer Unterricht.

Der gemeinsame rote Faden durch die verschiedenen Beiträge ist die Konzeptualisierung und Operationalisierung der Erklärkompetenz. Sie bildet in den verschiedenen Teilprojekten die Grundlage, um die spezifischen Fragestellungen zu bearbeiten und ermöglicht einen Vergleich der unterschiedlichen Ergebnisse. Die Besonderheit bei allen drei zugehörigen FALKE-Forschungsprojekten ist das parallele theoretische, konzeptuelle, methodische und administrative Vorgehen der beteiligten Unterrichtsfächer.

Beiträge des Symposiums

Zentrale Ergebnisse von FALKE-q: Fachspezifische Lehrkraftkompetenzen im Erklären – Qualität von Erklärungen

Stefan Krauss, Anita Schilcher, Alfred Lindl, Sven Hilbert
Universität Regensburg, Deutschland

Im ersten Projekt des FALKE-Forschungsprogramms (didaktische Qualität unterrichtlicher Erklärungen: FALKE-q) waren insgesamt 14 verschiedene Disziplinen an der Universität Regensburg beteiligt. Hierzu gehören neben den Methoden der empirischen Bildungsforschung die Unterrichtsfächer Biologie, Chemie, Deutsch, Englisch, Evangelische Religionslehre, Geschichte, Politische Bildung (in der Primarstufe), Kunst, Mathematik, Musik und Physik, sowie die deutsche Sprach- und die Sprechwissenschaft. Im Vortrag sollen die aus dem Projekt entstandene Buchpublikation sowie zentrale Ergebnisse (Schilcher et al., im Druck) im Überblick dargestellt werden.

Als zentrale Facette des fachdidaktischen Wissens wird Erklärkompetenz zwar bereits in Shulmans (1986) einflussreicher Taxonomie des Lehrerprofessionswissens beschrieben und demgemäß u. a. in der COACTIV-Studie (Kunter et al., 2011) und im Forschungsprojekt FALKO (Krauss et al., 2017) modelliert. Die Befundlage zu Merkmalen guten Erklärens ist allerdings je nach Bezugsdisziplin bislang noch uneinheitlich und insgesamt als defizitär zu erachten. Dieses Desiderat greift das FALKE-Forschungsprogramm auf.

Die Besonderheit bei allen drei zugehörigen FALKE-Forschungsprojekten ist dabei das parallele theoretische, konzeptuelle, methodische und administrative Vorgehen der beteiligten Unterrichtsfächer (Schilcher et al., 2021). In FALKE-q implementierten alle elf Fachdidaktiken einen im Wesentlichen identischen Onlinefragebogen, bei dem von insgesamt vier Teilstichproben (Schüler*innen, Lehrkräfte, Lehramtsstudierende, Fachdidaktiker*innen) die didaktische Qualität etwa dreiminütiger videografierter Erklärungen beurteilt werden musste. In einem ersten Durchlauf der Videos (sechs pro Fach) konnte für jedes Video – noch ohne Beeinflussung durch vorgegebene Kriterien – zunächst eine Gesamtnote vergeben werden (holistisches Globalurteil inkl. +/- Tendenz und offener Begründung). Anschließend wurden für einen zweiten Durchlauf der Videos in jedem Fach dieselben Konstrukte als Qualitätskriterien guter Erklärungen operationalisiert (genauer: Strukturiertheit und sprachliche Verständlichkeit jeweils mit vier Items, Adressatenorientierung mit fünf und der Sprech- und Köperausdruck mit sieben Items; als Kovariate wurde außerdem mit drei Items die Persönlichkeitswirkung des Erklärers/der Erklärerin erfasst; alle Items hatten ein sechsstufiges geschlossenes Antwortformat). Außerdem konnten pro Video noch zusätzliche fach- bzw. themenspezifische Aspekte bzw. Items implementiert werden.

Die untersuchte Gesamtstichprobe (in elf Unterrichtsfächern) bestand aus N = 3010 Personen aus vier verschiedenen Statusgruppen, die alle freiwillig sowie unentgeltlich an der Studie teilnahmen: Schüler*innen (N = 1396), Studierende (N = 833), Lehrkräfte (N = 442) sowie Didaktiker*innen (N = 339). Eine Triangulation der jeweiligen Blickwinkel bietet sich hier an, um Gemeinsamkeiten und Unterschiede in der Einschätzung guten Erklärens sowie diesbezügliche Einflussgrößen zu analysieren (so können Fachdidaktiker*innen z. B. als Expert*innen für theoretische Aspekte gesehen werden, während die Perspektive von Schüler*innen als die eigentlichen Adressat*innen unterrichtlicher Erklärungen von Interesse ist). Der eingesetzte Fragebogen weist in jeder Domäne und bezüglich aller Skalen zufriedenstellende interne Konsistenzen auf. In mehrfaktoriellen

Varianzanalysen mit gemischtem Design zeigen sich in allen Disziplinen überwiegend erwartungskonforme signifikante Differenzen zwischen den vier Statusgruppen sowie auch im Hinblick auf die Intentionen der fachdidaktisch jeweils in den Videos implementierten Manipulationen der Erklärungen (Schilcher et al., im Druck). Auch wenn die Prädiktoren Strukturiertheit, Adressatenorientierung, Sprech- und Körperausdruck sowie sprachliche Verständlichkeit fach- und statusgruppenabhängig variabel mit der Globaleinschätzung einer Erklärung assoziiert sind, erweisen sich in weiterführende Mehrebenenanalysen (z. B. Lindl et al., 2020) vor allem Adressatenorientierung, Strukturiertheit sowie Sprech- und Körperausdruck (nicht aber rein sprachliche Aspekte) transdisziplinär als signifikant und bedeutsam für die von den Studienteilnehmer*innen wahrgenommene globale Erklärqualität.

„Hübsch, sympathisch, nett – gut erklärt!“? Beurteilung von Erklärungen durch Schüler:innen und ihr Zusammenhang mit deren Leistung

Christina Knott, Lukas Lemberger, Stefanie Pietsch, Anita Schilcher
Universität Regensburg, Deutschland

Aus Sicht von Schüler:innen ist das Erklärenkönnen eine der zentralen Fähigkeiten einer Lehrkraft. Das Verstehen von Inhalten ist für sie der Sinn des schulischen Lernens (Wragg & Wood, 1984, Wörn, 2014). Eine gelungene Erklärung generiert nicht nur einen Wissens- und Kompetenzzuwachs bei den Schüler:innen sondern die von diesen zugesprochene Erklärkompetenz einer Lehrkraft beeinflusst maßgeblich, wie diese von ihnen wahrgenommen wird (Leisen, 2007).

Damit kann Erklären kann als eine zentrale fachdidaktische Kompetenz von Lehrkräften angesehen werden. Um eine hohe Qualität beim Erklären zu erreichen, müssen psychologisches, pädagogisches und fachliches Wissen integriert werden. Ein pragma-linguistischer Blick zeigt überdies die Komplexität des entsprechenden Sprachhandlungsmusters, das sowohl der fachlichen Struktur wie auch die Adressatenorientierung gerecht werden muss (Klein, 2009, Neumeister & Vogt, 2015). Dennoch gehört die explizite Professionalisierung im Erklären noch nicht zum Repertoire der universitären Ausbildung von Lehrkräften (Schilcher et al., 2017), dabei ließe sich Erklären im Hochschulsetting einfach und ohne großen Aufwand mit Studierenden als peers trainieren (Asen-Molz et al., 2022).

Der Beitrag stellt anhand der Fächer Deutsch und Sport das interdisziplinär entwickelte Seminarkonzept zur Schulung von Erklärkompetenz als Teilprojekt des Regensburger Forschungsprogramms FALKE-e (Fachspezifische Lehrkräftekompetenzen im Erklären - Expertise) vor. Er präsentiert Ergebnisse zur Bewertung von Erklärungen durch Schüler:innen und dem Zusammenhang mit deren Lernzuwachs.

FRAGESTELLUNGEN:

Welche Subfacetten der Erklärkompetenz (Fachlichkeit, Strukturiertheit, Adressatenorientierung, Sprech- und Körperausdruck, sprachliche Verständlichkeit, Persönlichkeit) hängen mit der Bewertung der Schüler:innen einer Unterrichtseinheit zusammen?

Hängen die Bewertung einer Erklärung durch Schüler:innen mit derem tatsächlichen Lernzuwachs zusammen?

METHODE

In einem fachdidaktischen Interventions-Seminar (2 SWS, ein Semester) werden den Lehramtsstudierenden verschiedene Subfacetten des Erklärens sowie zugehörige Strategien vermittelt und eingeübt. Nach einer einleitenden Theoriephase im Seminar erarbeiten die Studierenden eine 20-minütige Unterrichtsminiatur zu einem thematischen Schwerpunkt, die mindestens zwei Erklärphasen beinhaltet. Diese Miniatur erproben sie praktisch in einer Schulklasse (8./9. Jgst.) und werden dabei videografiert.

Die Schüler:innen der jeweiligen Klasse bewerten die Erklärungen sowie das Auftreten der Lehrkraft mithilfe eines Fragebogens, der die verschiedenen Subfacetten des Erklärens abbildet. Zusätzlich füllen sie vor und nach der Miniatur einen Kompetenztest (Paper-Pencil-Test, Prä-Post-Design) zu dem behandelten Thema aus. Durch diese Kompetenztests kann der Leistungszuwachs durch die Erklärungen ermittelt werden. Im Fach Deutsch werden die Kompetenztests mit einem aus der Theorie und dem Datenmaterial entwickelten Raster im Doppelblindverfahren codiert (N=200). Im Fach Sport wird der Kompetenzzuwachs mittels eines Wissenstests (Top-Down-Ansatz, gebundene & offene Aufgaben) durch die durchschnittliche Differenz der Punktzahlen in Prä- und Posttest ermittelt (N=270). Im Anschluss wird der Leistungszuwachs den Bewertungen der Erklärungen durch die Schüler:innen gegenübergestellt.

ERGEBNISSE UND IHRE BEDEUTUNG

Die Datenerhebung ist nahezu vollständig abgeschlossen. Die bisher erhobenen Daten lassen erkennen, dass für die Bewertung der Erklärung das Auftreten der Lehrkraft und die ihr entgegengebrachte Sympathie eine große Rolle spielen. Gleichzeitig beeinflusst eine positive Bewertung von Sprech- und Körperausdruck die Erklärwirkung maßgeblich.

Der tatsächliche Leistungszuwachs der Schüler:innen hängt nur mit verschiedenen Subfacetten des Erklärens, nicht aber mit der Bewertung der Erklärung zusammen. Besonders zeigt sich dies an der Subfacette Strukturiertheit.

In einem nächsten Schritt werden die Erklärungen von Expert:innen geratet, dadurch kann ermittelt werden, ob Expert:innen lernwirksame Erklärungen besser erkennen als Schüler:innen.

Adaptive Erklärkompetenz und Ungewissheitstoleranz von Lehramtsstudierenden

Franziska Schißlbauer, Stefanie Reimer, Nathalie Stegmüller, Arne Dittmer, Oliver Tepner, Stefan Krauss
Universität Regensburg, Deutschland

Im Vortrag sollen im Rahmen des Projekts FALKE-e erhaltene Befunde zur adaptiven, reflektierten Erklärkompetenz von Lehramtsstudierenden sowie deren Zusammenhang zur selbstberichteten Ungewissheitstoleranz und zur Selbstwirksamkeitserwartung vorgestellt werden.

Das Vorläuferprojekt FALKE-q (Fachspezifische Lehrkräftekompetenzen im Erklären) befasste sich mit fächerübergreifenden Qualitätskriterien guten Erklärens (Schilcher et. al, i.V.). Die Untersuchung adaptiver Erklärkompetenz im unterrichtlichen Setting stellt noch ein Desiderat dar. Unter Adaption wird dabei eine Maßnahme der Lehrkraft als Reaktion auf unvorhergesehene Schülerbeiträge und auf veränderte Situationen sowie eine Abweichung vom Unterrichtsplan verstanden (Parsons, 2012). Dabei wird zwischen Adaption “in der Planung“ und „in der Handlung“ unterschieden (Beck et al., 2008).

Adaption in der Planung erfolgt dabei idealerweise durch nachträgliche Reflexion. Donald Schön (1983) bezeichnete diese Fähigkeit auch als ‚reflection-on-action'. Er unterschied diese von der während des Unterrichts stattfindenden Reflexion (‚reflection-in-action'). Laut der Ständigen Kultusministerkonferenz spielt die Reflexion eine wichtige Rolle bei der Weiterentwicklung von professionellen Kompetenzen von Lehrkräften (Kultusministerkonferenz, 2014).

Eine treibende Kraft der Kompetenzentwicklung bei Lehrkräften ist, die Ungewissheit im Unterrichtsprozess als fruchtbares Moment und nicht als Problem zu akzeptieren (Keller-Schneider, 2018). Um der Ungewissheit von Unterricht gerecht zu werden, sind Adaptionen im unterrichtlichen Setting notwendig (Keller-Schneider, 2018). Unter Ungewissheitstoleranz versteht man nach Bonnet et al. (2021) ein Personenmerkmal, welches interindividuelle Unterschiede in der Bewertung ungewisser Situationen beschreibt (König & Dalbert, 2007).

Unter Selbstwirksamkeitserwartung versteht man die subjektive Gewissheit, eine neue oder schwierige Aufgabe auch dann erfolgreich bearbeiten zu können, wenn sich Widerstände in den Weg stellen (Schmitz & Schwarzer, 2000). Dabei kann die Selbstwirksamkeitserwartung sowohl das Planungs- als auch Unterrichtshandeln von Lehrkräften beeinflussen (Tschannen-Moran & Woolfolk Hoy, 2001). Diese kann durch mehrfaches Erklären ebenfalls verändert werden.

FRAGESTELLUNG

Lassen sich die Selbstwirksamkeitserwartungen und die Ungewissheitstoleranz von Lehramtsstudierenden bezüglich ihrer Erklärkompetenz im Rahmen eines universitären Seminars steigern, in dem die Studierenden Gelegenheit haben, eine ausgearbeitete (und sukzessiv verbesserte) Erklärung mehrfach an Schülergruppen zu geben? Hängen die untersuchten Konstrukte mit der adaptiven Erklärkompetenz der Studierenden zusammen?

METHODE

Im hier vorgestellten Verbundprojekt FALKE-e wurde unter anderem in den Fächern Biologie, Chemie und Mathematik je ein universitäres Didaktik-Seminar entwickelt, in dem die adaptive Erklärkompetenz von Lehramtsstudierenden anhand schriftlich ausgearbeiteter Planungen und Unterrichtsvideos an mehreren Messzeitpunkten gemessen wurde. Dabei gibt es jeweils eine „Erklärgruppe" und eine „Beobachtergruppe", wobei alle Studierenden eine bestimmte zwanzigminütige Erkläreinheit schriftlich planen müssen. Die Erklärgruppe (TG1) hält diese anschließend vor Schüler:innen, das Geschehen wird videografiert. Die Videos werden danach reflektiert (Selbstreflexion), überarbeitet und erneut vor einer anderen Schulklasse gehalten. Dieser Ablauf wird beim zweiten (und dritten) Messzeitpunkt wiederholt. Die Beobachtergruppe (TG2) plant ebenfalls eine zwanzigminütige Erkläreinheit, beobachtet die adaptiven Erkläreinheiten der TG1, reflektiert im Anschluss das Video der TG1 (Fremdreflexion) und überarbeitet im Nachgang die eigene Planung. Beobachtende und jeweilige Erklärende planen dabei das gleiche Unterrichtsthema.

Um die adaptive Erklärkompetenz zu messen, werden Kodiermanuale entwickelt. Bei TG1 können sowohl die schriftlichen Planungen als auch die Unterrichtsvideografien kodiert werden, bei TG2 nur die ausformulierten Planungen. Um die Selbstwirksamkeitserwartung und Ungewissheitstoleranz

sowie diesbezügliche Änderungen im Laufe des Seminars zu messen, werden beide Konstrukte im Pre-Post-Design erhoben.

VORLÄUFIGE ERGEBNISSE

Im Sommersemester 2022 wurde die erste Erhebung der Hauptstudie durchgeführt (mit über alle drei Fächer insgesamt n = 51 Lehramtsstudierenden). Es konnte bereits ein signifikanter Zuwachs (p=.001, t=7.93, df=27) der Einschätzung der Selbstwirksamkeitserwartung (□=.88) mit einer Effektstärke von dz=1.50 nachgewiesen werden. Bei der Ungewissheitstoleranz (□=.70) konnte bislang kein signifikanter Zuwachs (p=.586, t=0.56, df=17) mit dz=0.131 gefunden werden. Im Wintersemester 2022/23 wird die zweite Haupterhebung stattfinden (erwartet: n(Chemie)=24; n(Biologie)=20; n(Mathe)=40). Im Vortrag werden mit den vollständigen Hauptstudiendaten die Änderung der Ungewissheitstoleranz und Selbstwirksamkeit sowie der Zusammenhang zur adaptiven, reflektierten Erklärkompetenz vorgestellt. Auch wird auf mögliche Interaktionen mit der Gruppenzugehörigkeit TG1 vs. TG2 eingegangen.

FALKE-digital. Die Wirksamkeit von Erklärvideos im Rahmen von Flipped-Classroom

Daria Podwika, Anita Schilcher
Universität Regensburg, Deutschland

Im Kontext der Corona-Pandemie, die zur Einführung des Distanzunterrichts führte, haben Erklärvideos als professionelles Vermittlungsformat Konjunktur. 45 % der Schülerinnen und Schüler geben in der gerade neu erschienenen Sonderausgabe der JIM-Studie, die sich im Jahr 2020 mit Lernen und Freizeit in der Corona-Krise beschäftigt, an, dass sie Tutorials im Internet zum Lernen heranziehen (vgl. mpfs, 2020). Dabei nimmt YouTube mit 83 % den ersten Platz bei der Nutzung medialer Inhalte ein. Dadurch, dass Erklärvideos zu einem bedeutsamen Bildungsmedium des 21. Jahrhunderts zählen (vgl. Sailer & Figas, 2015), rücken nicht nur die Produktion guter Erklärvideos (vgl. Kulgemeyer, 2020)), sondern auch die Rezeption von Erklärvideos zunehmend in den Fokus der Forschung. Dass der Einsatz von Erklärvideos nicht per se zu positiven bzw. erfolgreichen Lernergebnissen führt, zeigen Wolf und Kulgemeyer (2020). Damit Lernende deren Inhalte nicht nur oberflächlich verarbeiten, sondern diese in gesichertes Wissen überführen, ist die aktive Steuerung von Verarbeitungsprozessen von besonders hoher Relevanz (vgl. Dube, 2021; vgl. Weidlich & Spannagel, 2014). Vor diesem Hintergrund sollen Erklärvideos nicht isoliert in das Unterrichtsgeschehen eingebettet werden, sondern Lernprozesse initiieren und unterstützen (Meller, 2017). Dabei spielen auch Lernstrategien für die zielgerichtete Rezeption der Erklärvideos eine zentrale Rolle für das selbstgesteuerte Lernen. Aufgabe von Unterricht ist es, diesen Prozess des selbstgesteuerten Lernens zu unterstützen und den Aufbau von Selbstlernkompetenzen zu ermöglichen (Schulz, 2020). Hierfür eignet sich besonders das Flipped-Classroom-Konzept. In diesem werden die zentralen Aktivitäten des Lehrens und Lernens umgekehrt. Zuhause sollen sich die Schülerinnen und Schüler die Inhalte individuell mithilfe von Erklärvideos aneignen, während der Präsenzunterricht stärker für die Vertiefung, Übung, Anwendung oder Reflexion des Gelernten genutzt wird (vgl. Ebel, 2018). Um auch in diesem Unterrichts-Setting positive Lerneffekte zu erzielen, sollen Erklärvideos nicht gesondert eingesetzt werden, sondern um eine aufgabenbasierte Vorbereitungsphase ergänzt werden (vgl. Weidlich & Spannagel, 2014).

Die Forschungsgruppe FALKE-d untersucht die Wirksamkeit von Erklärvideos im Rahmen eines Flipped-Classroom-Settings in unterschiedlichen Fächern. Am Projekt nehmen fünf Fachdidaktiken (Chemie, Mathematik, Deutsch, Musik, Grundschulpädagogik) teil. Während es im mathematischen und naturwissenschaftlichen Bereich bereits viele gut etablierte Erklärvideo-Plattformen gibt, finden sich etwa im Fach Deutsch – etwa beim literarischen Lernen – kaum Angebote (vgl. Schilcher et al., 2022). Deshalb wurde für das Fach Deutsch im Rahmen des Projekts eine Lernplattform entwickelt, auf der Schülerinnen und Schüler mit Hilfe von Erklärvideos lernen, wie sie literarische Texte und Kurzfilme analysieren und interpretieren (Schilcher & Pissarek, 2018). Eine Treatmentgruppe erhält hierzu spezifische Rezeptionsstrategien, um mit den Videos zu arbeiten. Eine zweite Gruppe arbeitet ohne Strategien mit den Erklärvideos. Die dritte Gruppe arbeitet mit den gleichen Materialien im regulären Unterricht (Erklärung im Unterricht, Übung zuhause).

FRAGESTELLUNGEN

Ist das Flipped-Classroom-Konzept für das literarische Lernen erfolgreicher als der reguläre Unterricht?

Profitieren die Kinder der Strategiegruppe durch die Vermittlung und Automatisierung spezifischer Rezeptionsstrategien?

METHODE

Anhand eines Prä-, Post-, Follow-Up-Designs mit zwei Treatmentgruppen und einer Kontrollgruppe wird der Leistungszuwachs anhand selbst erstellter Kompetenztests bei den Schülerinnen und Schülern gemessen. Mit Hilfe von Tracking-Daten kann das Nutzungsverhalten der Schülerinnen und Schüler im Umgang mit den Erklärvideos nachvollzogen werden und mit den Leistungsdaten in Zusammenhang gebracht werden. Pilotierungsdaten zeigen, dass bereits Grundschülerinnen und Grundschüler (n = 40) mit dem Flipped-Classroom-Konzept zurechtkommen und Lernzuwächse erzielen.

ERGEBNISSE

Die Haupterhebung findet im November und Dezember 2022 in mindestens 25 Klassen der vierten bis sechsten Jahrgangsstufe statt. Im Rahmen des Vortrags werden erste Ergebnisse vorgestellt.

H-15: Unterrichtsplanungskompetenz angehender Lehrkräfte: Planungsentscheidungen erfassen, analysieren und reflektieren

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* T03 R03 D75

Unterrichtsplanungskompetenz angehender Lehrkräfte: Planungsentscheidungen erfassen, analysieren und reflektieren

Chair(s): **Thomas Rey** (Pädagogische Hochschule St.Gallen, Schweiz), **Valentin Unger** (Pädagogische Hochschule St.Gallen), **Hendrik Lohse-Bossenz** (Universität Greifswald)
*Diskutant*in(nen):* **Christian Brühwiler** (Pädagogische Hochschule St.Gallen)

Die Unterrichtsplanung gilt als zentrale Aufgabe des Lehrberufs (Jäger & Maier, 2019) und beeinflusst die Unterrichtsqualität (Gassmann, 2013), weshalb der Aufbau von Unterrichtsplanungskompetenz zu den wesentlichen Qualifizierungsaufgaben der Lehrpersonenbildung zählt (Schnebel et al., 2017). Als empirischer Forschungsgegenstand ist diese Kompetenz dagegen nach wie vor auffällig unterrepräsentiert (König & Rothland, 2022).

Darum soll im Symposium anhand aktueller empirischer Befunde, in die sowohl allgemein- als auch fachdidaktische Perspektiven einfließen, genauer ergründetet werden, wie und welche unterrichtlichen Planungsentscheidungen angehende Lehrkräfte treffen. Gleichzeitig werden Einblicke in die jeweiligen Messverfahren geboten, um auch die methodische Diskussion der empirischen Unterrichtsplanungsforschung zu befördern.

Im Grundlagenbeitrag 1 wird vorgeschlagen, dass die Planungsprobleme von angehenden Lehrpersonen durch die kognitiv überlastende hohe Interaktivität der Planungselemente beim interdependenten Unterrichtsplanungsentscheiden erklärt werden können. Für erste Belege dieser Erklärung wird mithilfe einer experimentellen Studie geprüft, ob die intrinsische Arbeitsgedächtnisbelastung in Abhängigkeit von der Menge der Planungselemente und ihrer Verknüpfung bei der Unterrichtsplanung zunimmt. Die vielversprechenden Ergebnisse eröffnen eine neue Perspektive auf die Forschung und Lehre zur Unterrichtsplanung.

Der fachdidaktische Beitrag 2 widmet sich der Rückkopplung diagnostischer und didaktischer Handlungen in der Planungs- und Realisierungsphase, um Unterricht an die Voraussetzungen der Lernenden anzupassen und erfolgreiches Lernen zu bewirken. Untersucht wird, welche Rückkopplungen im Sportunterricht bestehen und inwiefern sich diese bei Sportlehrkräften mit unterschiedlicher diagnostischer Kompetenz ähneln bzw. unterscheiden. Die qualitativen Ergebnisse zeigen, dass Rückkopplungen seltener in der Planungs- als in der Realisierungsphase vorhanden sind. Unterschiede zeigen sich zwischen Sportlehrkräften u.a., inwiefern sich diagnostische und didaktische Handlungen im Unterricht überlagern oder aufeinander abgestimmt werden.

Die nachfolgenden Beiträge betrachten das adaptive Planen für heterogene Lerngruppen. Zunächst werden im Beitrag 3 Ergebnisse aus einer qualitativen Studie zur Unterrichtsplanung im Vorbereitungsdienst Deutsch vorgestellt. Der Beitrag geht der Frage nach, welchen Stellenwert Diagnostik, Differenzierung und deren adaptive Verknüpfung in schriftlichen Unterrichtsplanungen haben. Die Ergebnisse zeigen, dass in den meisten Unterrichtsplanungen nur in einem schwachen bis mittleren Grad adaptiv geplant wird und werfen in ihrer Gesamtheit Fragen zur Gestaltung der ersten und zweiten Ausbildungsphase auf.

Daran anknüpfend, fokussiert der abschließende Beitrag 4 adaptive Planungsentscheidungen von Referendar*innen aus einer allgemeindidaktischen Perspektive. Basierend auf einem theoretischen Strategie-Kategorien-Modell adaptiver Planungskompetenz wurde zur Erfassung von Planungsentscheidungen im Umgang mit Heterogenität in schriftlichen Unterrichtsplanungen das Messinstrument *AIPako* entwickelt. Anhand allgemeindidaktischer Kriterien zum Diagnostizieren, Differenzieren und zur Klassenführung wurden schriftliche Unterrichtsplanungsentscheidungen extrahiert und quantifiziert. Die Ergebnisse validieren das Testinstrument unter Kontrolle schriftsprachlicher Oberflächenmerkmale, konstatieren das postulierte Modell und identifizieren Primar- und Stützstrategien im adaptiven Planen.

Beiträge des Symposiums

Scheitern Anforderungen an die Unterrichtsplanung an den Grenzen der Informationsverarbeitung? Eine experimentelle Studie zur intrinsischen Arbeitsgedächtnisbelastung von angehenden Lehrkräften beim Planen

Simon Küth, Daniel Scholl
Universität Siegen

Theoretischer Hintergrund und Fragestellung. Wegen der strukturellen Ungewissheit von Unterricht (Bonnet et al., 2021) sollten sich (angehende) Lehrkräfte „nicht der Illusion hingeben, mit der Planung vollständige Sicherheit für die eigentliche Unterrichtshandlung zu gewinnen“ (König & Rothland, 2022, S. 804). Dennoch wird in der gedanklichen Vorwegnahme des künftigen Unterrichts eine so große Chance für seine kontinuierliche Qualitätsentwicklung gesehen (Wengert, 1989), dass die Planungskompetenz bereits ab dem Studium gefördert werden soll (KMK, 2019) und hohe Ansprüche an das zugrundeliegende (allgemein)didaktische Planungs- und Konzeptionswissen (Baumert & Kunter, 2006) gestellt werden: Unter anderem im Sinne des Interdependenzprinzips (Vogelsang & Riese, 2017), das in didaktischen Ansätzen (Scholl, 2018), im Decision Making-Paradigma (z.B. Shavelson & Stern, 1981) und in aktuellen generischen (vgl. Rothland, 2021) und fachspezifischen Kompetenzmodellen (z.B. Kirsch, 2020; Schröder et al., 2020) gleichermaßen vertreten wird, sollen (angehende) Lehrkräfte ihre inhaltlichen, methodischen, thematischen und medialen Planungsentscheidungen kohärent untereinander und mit den Lernvoraussetzungen verknüpfen.

In der Planungspraxis scheitern insbesondere Planungsanfänger*innen mit ihrer lediglich linearen Planung (Westerman, 1991) in isolierten Planungsbereichen (Bromme, 1981) allerdings so häufig an einer solchen Verknüpfung (z.B. Gassmann, 2013), dass es überrascht, dass bisher kaum Erklärungen für dieses Missverhältnis zwischen Anspruch und Verwirklichung vorliegen. Deshalb wird im vorgeschlagenen Beitrag aus der Perspektive der Cognitive Load-Theorie (Sweller et al., 2019) untersucht, ob das interdependente Planungsentscheiden als komplexe Aufgabe mit einer hohen Elementinteraktivität (Küth et al., 2021; Schrader & Schöb, 2016) durch die Möglichkeiten der Informationsverarbeitung begrenzt wird: Diese Aufgabe könnte vor allem dann mit einer hohen intrinsischen Belastung des kapazitätsbegrenzten Arbeitsgedächtnisses einhergehen, wenn – wie bei Planungsanfänger*innen – noch keine arbeitsgedächtnisentlastenden Planungsroutinen vorhanden sind (Sweller & Chandler, 1994). Um diesen Erklärungsansatz zu prüfen, wurde in einer experimentellen Studie im Within-Design gefragt, inwiefern die intrinsische Belastung in Abhängigkeit von der Menge der Planungselemente und ihrer Verknüpfung zunimmt.

Methode. N = 98 Lehramtsstudierenden wurden in einem Onlinetest in randomisierter Reihenfolge vier geschlossene Planungsvignetten mit jeweils drei mal vier Planungsentscheidungen, von denen laut Expert*inneneinschätzung jeweils eine die kohärenteste darstellt, auf unterschiedlichen Komplexitätsstufen vorgelegt. Variiert wurden diese Stufen über die Menge der jeweils verbundenen Planungselemente (Stufe 1: 2, Stufe 2: 3 und Stufe 3: 4 bis 5 Elemente). Die Studierenden wurden direkt im Anschluss an jede Aufgabenbearbeitung um ihre Einschätzung der Arbeitsgedächtnisbelastung mittels der validen und reliablen Items von Klepsch et al. (2017) zur intrinsischen und extrinsischen kognitiven Belastung gebeten. Die extrinsische Belastung wurde als etwaige Folge der zwischen den Komplexitätsstufen variierenden Vignettenlänge kontrolliert.

Ergebnisse. Die Ergebnisse stützen die Annahme, dass die intrinsische Arbeitsgedächtnisbelastung je nach Anzahl der zu berücksichtigenden Planungselemente unterschiedlich hoch ist: Die durchschnittliche Einschätzung der intrinsischen kognitiven Belastung nimmt über die Komplexitätsstufen zu (MK1 = 3,15, SD K1 = 1,04; M K2 = 3,44, SD K2 = 1,04; M K3 = 3,67, SD K3 = 1,10). Diese Unterschiede sind für jedes Paar signifikant ($F(1{,}69, 164{,}17) = 33{,}11$, $p < 0{,}001$). Für die extrinsische Belastung zeigt sich ein Unterschied nur für den paarweisen Vergleich der Komplexitätsstufen 1 und 3 (MK1 = 2,81, SD K1 = 0,91; M K2 = 2,91, SD K2 = 0,93; M K3 = 2,96, SD K3 = 0,92; $F(1{,}95, 188{,}84) = 3{,}87$, $p = 0{,}24$). Die intrinsische Belastung wird für jede Komplexitätsstufe höher eingeschätzt als die extrinsische Belastung.

Diskussion. Die Ergebnisse dieser Grundlagenstudie sind ein erster Hinweis auf die Erklärung der Planungsschwierigkeiten von Planungsanfänger*innen durch eine hohe intrinsische Arbeitsgedächtnisbelastung. Da die Belastung lediglich über Selbsteinschätzungen und vorgegebene Planungsentscheidungen erfasst wurde, sind Folgestudien notwendig. Dennoch lassen sich aus den Ergebnissen vorsichtige Schlüsse über neue Ansatzpunkte für die Förderung der Planungskompetenz im Lehramtstudium ziehen, für die sich beispielsweise eine Reduktion der Elementinteraktivität in den ersten Planungsversuchen, z.B. mittels isolated-elements procedures, (Pollock et al., 2002) anbieten könnte.

Rückkopplung diagnostischer und didaktischer Handlungen in der Planungs- und Realisierungsphase von Sportunterricht

Anneke Langer
Europa-Universität Flensburg

Einleitung. Eine der zentralen Herausforderungen für Grundschullehrkräfte ist der Umgang mit heterogenen Lernvoraussetzungen von Kindern. Dafür bedarf es der Wahrnehmung und Beurteilung individueller Unterschiede - einer pädagogischen Diagnose (Behrmann & van Ophuysen, 2017). Akkurate pädagogische Diagnosen zählen zu den professionellen Aufgaben von Lehrkräften und ermöglichen es, begründete pädagogische Entscheidungen zu treffen (Baumert & Kunter, 2006). Die zugrundeliegende Fähigkeit wird als diagnostische Kompetenz bezeichnet und ist fachspezifisch ausgebildet (Schrader, 2011). Um Sportunterricht angemessen und lernförderlich an die Lernvoraussetzungen (z.B. Lernmotivation, Motorische Kompetenzen) von Kindern anzupassen, ist eine enge Verknüpfung diagnostischer und didaktischer Handlungen erforderlich (Seyda & Langer, 2020), aber bisher kaum erforscht. Hier setzt der Vortrag an und geht erstens der Frage nach, welche

Rückkopplungen diagnostischer und didaktischer Handlungen in der Planungs- und Realisierungsphase von Sportunterricht bestehen. Zweitens wird untersucht, inwiefern bei Sportlehrkräften mit unterschiedlicher diagnostischer Kompetenz Gemeinsamkeiten oder Unterschiede in der Rückkopplung vorhanden sind. Das Wissen darüber ermöglicht es, Aspekte einer guten Diagnose als Voraussetzung für eine individuelle Förderung zu identifizieren.

Methode. Mit zehn Grundschulsportlehrkräften mit unterschiedlicher diagnostischer Kompetenz wurden Leitfadeninterviews zu ihrem diagnostischen und didaktischen Handeln im Sportunterricht geführt. Damit sind die Interviewaussagen vergleichbar. Die diagnostische Kompetenz wurde mittels Rangkomponente berechnet (Schrader, 2011). Die Auswertung erfolgte erstens mittels strukturierender Inhaltsanalyse (Mayring, 2015) zu den Kategorien „Rückkopplung in der Planungsphase“ (z.B. Berücksichtigung der Lernvoraussetzungen und Hilfestellungen planen) und „Rückkopplung in der Realisierungsphase“ (z.B. Lernprozess einschätzen, Rückmeldungen geben und auf Umsetzungsschwierigkeiten reagieren). Zweitens wurden drei Gruppen (hohe, mittlere, niedrige diagnostische Kompetenz) in Anlehnung an die Rangkomponenten gebildet. Bei der Auswertung wurden bei den drei Gruppen mittels qualitativen Gruppenvergleich (Rädiker & Kuckartz, 2019) Gemeinsamkeiten und Unterschiede zu den o.g. Kategorien untersucht.

Ergebnisse und Diskussion. In der Planungsphase äußern sich nur wenige Sportlehrkräfte über Rückkopplungen diagnostischer und didaktischer Handlungen. Die Überlegungen sind u.a. abhängig vom Inhaltsbereich. Während z.B. beim Turnen Hilfestellungen und Differenzierungen für unterschiedliches motorisches Können überlegt werden, erfolgen solche Planungen nicht für das Ringen und Kämpfen. Darüber hinaus werden z.B. Differenzierungsmaßnahmen für sportmotorisch stärkere und schwächere Kinder geplant. Allerdings wird nur selten thematisiert, ob und welche didaktischen Maßnahmen für Kinder mit durchschnittlichen sportmotorischen Kompetenzen geplant werden. In der Realisierungsphase äußern sich alle Sportlehrkräfte zu Rückkopplungen diagnostischer und didaktischer Handlungen. Sie zeigen sich u.a. darin, dass Differenzierungen bei Stationsarbeit oder Übungen durchgeführt werden, um auf die heterogenen Lernvoraussetzungen der Kinder einzugehen. Wenige Sportlehrkräfte äußern, dass sie didaktische Maßnahmen in der Realisierungsphase nutzen, um diagnostizieren zu können. Beispielsweise führen wenige Sportlehrkräfte bestimmte Spiele durch, bei denen sie die Anstrengungsbereitschaft und Motivation der Kinder beobachten und beurteilen können.

Der qualitative Gruppenvergleich ergibt, dass überwiegend Sportlehrkräfte mit hoher diagnostischer Kompetenz versuchen, jedes Kind individuell zu beurteilen. Zudem schaffen sie sich durch ihre didaktischen Handlungen Zeit zum Diagnostizieren. Hingegen schätzen Sportlehrkräfte mit mittlerer und niedriger diagnostischer Kompetenz die Lernvoraussetzungen von Gruppen oder der gesamten Klasse ein. Bei ihnen überlagern sich oft diagnostische und didaktische Handlungen, wodurch Beurteilungsverzerrungen entstehen können.

Neben der Diagnosegenauigkeit können für ein lernförderliches didaktisches Setting auch andere Kriterien bedeutend sein, wie z.B. eine leichte Überschätzung im Sinne von Vygotsky’s zone of proximal development. Bei den Ergebnissen ist u.a. zu diskutieren, inwiefern die Sportlehrkräften in

eine Überforderungsfalle hineingeraten (Heymann, 2010). Diese kann entstehen, wenn die Sportlehrkraft jedem Kind ein individuell passgenaues Lernangebot machen soll. Im Vortrag werden neben der Diagnosegenauigkeit weitere Aspekte für ein lernförderliches didaktisches Setting ebenso diskutiert wie das methodische Vorgehen und die Limitationen der Studie.

Der Stellenwert von Diagnostik und Differenzierung in schriftlichen Unterrichtsplanungen im Fach Deutsch

Ilka Fladung
Universität zu Köln

Theoretischer Hintergrund. Angehende Deutschlehrkräfte sollen während des Studiums und des Vorbereitungsdienstes u. a. die Planung, Durchführung und Analyse von Unterricht in heterogenen Lerngruppen erlernen (KMK, 2008). Dass dafür eine differenzierende Unterrichtsgestaltung eine zentrale Rolle spielt, wird dabei nicht nur in Lehramtscurricula auf unterschiedlichen Ebenen festgehalten (bspw. MSW NRW, 2011), sondern auch in der Deutschdidaktik meist als Selbstverständlichkeit gehandelt (u.a. von Brand & Brandl, 2017; Betzel et al., 2020; Hennies & Ritter, 2016). Empirisch abgesicherte Aussagen dazu, ob und in welcher Form angehende Deutschlehrkräfte heterogenen Lernvoraussetzungen tatsächlich mit der Planung differenzierender Unterrichtssettings begegnen, liegen bislang jedoch nicht vor.

An dieser Forschungslücke setzte die Studie an, fokussierte jedoch im Speziellen den adaptiven Einsatz von Differenzierungsmaßnahmen, da bereits belegt werden konnte, dass Differenzierung zwar nicht ‚per se' wirksamer ist als andere Formen der Unterrichtsgestaltung (Kulik & Kulik, 1982; Lou et al., 1996), dem adaptivem Unterricht hingegen „konsistent mittlere positive Effekte auf kognitive Lernergebnisse" (Brühwiler, 2014: 71) zugeschrieben werden dürfen – und zwar vor allem in leistungsheterogenen Klassen (Beck et al., 2008). Wird Unterricht adaptiv geplant, dann soll gezielt eine Passung von Lernvoraussetzungen und Lernumgebungen hergestellt werden (Corno & Snow, 1986; Dumont, 2019; Martschinke, 2015; Wang, 1980). Sollen dafür Lernvoraussetzungen wie bspw. der Vorwissensstand im Hinblick auf einen konkreten Unterrichtsgegenstand, die allgemeine Leistungsfähigkeit, motivationale und metakognitive Strategien oder motivationale Voraussetzungen (Gräsel, Decristan & König, 2017) festgestellt werden, so sind diagnostische Urteile (formeller oder informeller Art) von Nöten (Schrader, 2013; Südkamp & Praetorius, 2017).

Fragestellungen. Der vorliegende Beitrag ergänzt bisherige Studien zur Differenzierung im Deutschunterricht (bspw. Letzel & Otto, 2019) sowie fächerübergreifende Studien zur Planung im Vorbereitungsdienst (bspw. König, Buchholtz & Dohmen, 2015; Weingarten, 2019) und lotet entlang der folgenden Forschungsfragen den Stellenwert von Diagnostik und Differenzierung in schriftlichen Unterrichtsplanungen im Fach Deutsch aus:

- FF 1: In welcher Form werden unterschiedliche Heterogenitätsdimensionen bei Urteilen über Lernvoraussetzungen in Planungen relevant gesetzt?

- FF 2: Wie häufig basieren die Informationen in den Lerngruppenbeschreibungen auf formellen oder informellen diagnostischen Urteilen?

- FF 3: Sind die zentralen Lernaufgaben in Planungen als differenzierte Aufgaben angelegt?

- FF 4: Wie groß ist der Anteil der Planungen, denen ein adaptiver, d. h. proaktiver Umgang mit Heterogenität im Sinne Weinerts (1997) attestiert werden kann?

Methode. Grundlage dieser Studie sind schriftliche Unterrichtsplanungen aus NRW und Berlin zu zwei Messzeitpunkten (T1: n=172; T2: n= 138) und begleitende Fragebogenerhebungen (s. auch König et al. 2017; 2020). Die Auswertung der schriftlichen Planungen erfolgte mittels Qualitativer Inhaltsanalyse. Das Kategoriensystem für die inhaltlich-strukturierende Analyse wurde in einem deduktiv-induktiven Mischverfahren erstellt. 55 % der Entwürfe wurden von der Autorin und einer geschulten Zweitkodiererin ausgewertet. Werte von k = .94 (T1) und k = .80 (T2) weisen auf eine sehr gute Interraterreliabilität hin.

Ergebnisse. Der Vortrag stellt inhaltsanalytische Ergebnisse zu allen FF vor und stellt ihnen interessante Angaben zu selbst eingeschätzter Kompetenz und Lerngelegenheiten zur Seite. Insgesamt legen die Befunde nahe, dass im Vorbereitungsdienst Deutsch die Diagnostik von Lernvoraussetzungen kaum dem entspricht, was seitens Fachdidaktik und Curricula als erwünscht gelten kann (FF1 und FF2). So sind bspw. in 15 % aller untersuchten Planungen ausschließlich gemeinsame Lernvoraussetzungen der Klasse Gegenstand der Betrachtung; über Unterrichtsbeobachtungen hinausgehend lassen sich kaum Erwähnungen von Handlungen finden, die üblicherweise unter das diagnostische Repertoire von Deutschlehrkräften gefasst würden, und über den Einsatz (förder-)diagnostischen Materials wird in keiner Planung berichtet. Für die Differenzierung hingegen ist die Kluft zwischen dem Gewünschten/Geforderten und dem Status quo zwar als kleiner einzuschätzen, da zu T1 ca. 60 % und zu T2 95 % aller Lernaufgaben differenzierte Aufgaben sind, jedoch werden diese Differenzierungsmaßnahmen nur in 53 % (T1) bzw. 39 % (T2) adaptiv, d. h. passend zu den dargestellten Lernvoraussetzungen, geplant (FF3 und FF4).

Allgemeindidaktische Kriterien zur Erfassung adaptiver Planungskompetenz im Umgang mit Heterogenität (AlPako): Validierung eines Testinstruments

Thomas Rey[1], Valentin Unger[1], Hendrik Lohse-Bossenz[2]
[1]Pädagogische Hochschule St.Gallen, [2]Universität Greifswald

Zusammenfassung. Bislang liegen kaum Befunde zu Unterrichtsplanungskompetenz bei angehenden Lehrkräften im Umgang mit heterogenen Lerngruppen in den einzelnen Phasen der Lehrer*innenbildung vor. In diesem Zusammenhang wird eine adaptive Planungskompetenz, d. h. die Expertise zur Passung von fachunspezifischen Voraussetzungen und Bedürfnissen der Lernenden mit planerischen Unterrichtsentscheidungen als generische Anforderung, in diesem Beitrag fokussiert. Mit Hilfe des neu entwickelten Messinstrumentes AlPako (Allgemeindidaktische Kriterien zur Erfassung adaptiver Planungskompetenz im Umgang mit Heterogenität) soll adaptive Planungskompetenz von Referendar*innen abgebildet werden. Hierzu werden schriftliche Unterrichtsentwürfe hinsichtlich allgemeindidaktischer Kriterien analysiert, um die darin getroffenen Planungsentscheidungen der angehenden Lehrkräfte zu quantifizieren. Die Ergebnisse identifizieren den adäquaten Umgang mit Heterogenität als querschnittliche Dimension professioneller Lehrkompetenz unter Kontrolle schriftsprachlicher Oberflächenmerkmale. Die konfirmatorischen Faktorenanalysen weisen eine gute Modellpassung auf und konstatieren ein Strategie-Kategorien-Modell adaptiver Planungskompetenz. Auf Grundlage der empirischen Befunde können vielfältige

Einsatzmöglichkeiten der allgemeindidaktischen Kriterien auf die schulische Praxisbegleitung von angehenden Lehrpersonen diskutiert werden.

Theoretischer Hintergrund. Während bisherige Studien zur professionellen Kompetenz von Lehrkräften eher die Unterrichtsdurchführung in den Blick nehmen, scheint die Fähigkeit von Lehrkräften Unterrichtsverläufe zu planen, bislang in der empirischen Modellbildung weniger akzentuiert worden zu sein (König & Rothland, 2022; Rothland, 2021; Wernke & Zierer, 2017). Die wenigen aktuellen Forschungsbefunde zu Unterrichtsplanungskompetenz bei angehenden Lehrkräften beziehen sich nahezu ausschließlich auf fachdidaktische Untersuchungen (z. B. Fladung, 2022; Kirsch, 2022; Riese et al., 2022). Im vorliegenden Beitrag wird adaptive Planungskompetenz aus einer allgemeindidaktischen Perspektive fokussiert, welche entlang dreier theoretisch begründeter Subfacetten des professionellen Handelns im Umgang mit Heterogenität konzeptualisiert wird: (1) In heterogenen Gruppen diagnostizieren können, (2) in heterogenen Gruppen differenzieren können und (3) Klassenführung in heterogenen Gruppen (Beck et al., 2010; Franz, Wacker & Heyl, 2018; Rey, 2022). Basierend auf einer Pilotierungsstudie (Rey, Lohse-Bossenz, Wacker & Heyl, 2018) wurden die spezifischen Kriterien zur Messung adaptiver Planungskompetenz im Umgang mit Heterogenität weiterentwickelt und ein theoretisches Rahmenmodell postuliert (Rey, 2022), welches Primär- und Stützstrategien beinhaltet.

Fragestellungen. Die standardisierte Messung adaptiver Planungskompetenz angehender Lehrkräfte im Umgang mit Heterogenität stellt in dieser Untersuchung das zentrale Forschungsvorhaben dar. Vor diesem Hintergrund liegen diesem Beitrag einerseits die Validierung und Dimensionalitätsprüfung der modifizierten Kriterien zur Erfassung adaptiver Planungskompetenz (AlPako) unter Berücksichtigung schriftsprachlicher Oberflächenmerkmale zugrunde. Zusammenhang geht er der Frage nach, ob die Erfassung des Konstrukts in schriftlichen Unterrichtsentwürfen abhängig von schriftsprachlichen Oberflächenmerkmalen ist. Andererseits gilt es das theoretisch postulierte Strategie-Kategorien-Modell empirisch zu prüfen.

Methode. Datengrundlage dieser Validierungsstudie bilden 82 schriftliche Unterrichtsplanungen aus 23 Fächern und Fächerverbünden von Lehramtsanwärter*innen der Sekundarstufe I in Baden-Württemberg zu Beginn ihres Vorbereitungsdienstes. Ziel ist es, die darin enthaltenen Planungs¬entscheidungen extrahieren und quantifizieren zu können. Als Übereinstimmungsmaß für den Kodiervorgang wird die Interrater-Reliabilität Cohens Kappa (κ) herangezogen. Ferner wird die Itemschwierigkeit der einzelnen Kriterien als auch die Dimensionalität der Datenstruktur sowie die Passung der einzelnen Kriterien überprüft. Zudem werden die schriftlichen Unterrichtsplanungen anhand schwierigkeitsgenerierender Faktoren auf der Textoberfläche des Lesbarkeitsindex (Lenhard & Lenhard, 2014-2017) klassifiziert als auch die Zeichenanzahl operationalisiert.

Ergebnisse. Die berechneten Kappa-Werte zur Überprüfung der Interrater-Reliabilität (M = 0.904) als auch die Modellfitindizes (χ^2 (N = 82, df = 34) = 36.077, p = .371, χ^2/df = 1.061, CFI = 0.953, RMSEA = .027, WRMR = 0.755) weisen zufriedenstellende Ergebnisse auf. Darüber hinaus konnte kein signifikanter Zusammenhang zwischen gemessener adaptiver Planungskompetenz und den schriftsprachlichen Oberflächenmerkmalen gezeigt werden. Die Analysekriterien des

Testinstrumentes werden demnach nicht von der sprachlichen Komplexität oder der Zeichenanzahl der geschriebenen Unterrichtsentwürfe beeinflusst. Auf Basis der Befunde hinsichtlich Dimensionalität der Datenstruktur konnte zudem das theoretisch postulierte Strategie-Kategorien-Modell empirisch bestätigt werden: Es konnten die Subfacetten (1 & 2) in heterogenen Gruppen diagnostizieren und differenzieren können als Primärstrategie und die Subfacette (3) Klassenführung in heterogenen Gruppen als Stützstrategie adaptiver Planungskompetenz im Umgang mit Heterogenität identifiziert werden.

H-16: Bewertung von Schüler*innentexten - von Menschen und Maschinen

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* T03 R03 D89

Bewertung von Schüler*innentexten - von Menschen und Maschinen

Chair(s): **Julian Lohmann** (Christian-Albrechts-Universität zu Kiel, Deutschland), **Frederike Strahl** (Christian-Albrechts-Universität zu Kiel, Deutschland)
*Diskutant*in(nen):* **Detlef Urhahne** (Universität Passau)

Schriftsprachliche Kompetenzen sind eine entscheidende Voraussetzung für schulischen, akademischen und beruflichen Erfolg (Crossley & McNamara, 2016) und somit als eine Schlüsselkompetenz für gesellschaftliche Teilhabe zu sehen (Kultusministerkonferenz, 2014). Daher ist es ein wichtiges Ziel, die Schreibkompetenzen von Schüler*innen im Unterricht zu fördern. Die akkurate Einschätzung des Leistungsstandes der Lernenden gilt dafür als eine wichtige Voraussetzung (Schrader, 2013). Die Urteilsgenauigkeit von Lehrkräften bei Textbeurteilungen wurde bisher allerdings kaum empirisch untersucht.

In den ersten beiden Studien dieses Symposiums wird die Urteilsgenauigkeit deshalb zum zentralen Gegenstand. Methodische Herausforderungen bei Studien zur Textbeurteilung sind zum einen der Aufwand, der mit Bewertungsprozessen verbunden ist und zum anderen die Unsicherheit bezüglich der Güte eines Textes. Die „wahre“ Qualität eines Textes festzustellen ist unmöglich. Als Maßstab müssen Urteile von geschulten Expert*innen herangezogen werden, die insbesondere bei der Bewertung analytischer Teilaspekte von Textqualität allerdings eine geringe Interraterübereinstimmung aufweisen können (z.B. Horbach et al., 2017)

Um zukünftig eine Alternative zur arbeitsintensiven menschlichen Bewertung von Schreibleistungen zu gewinnen, werden derzeit die Potenziale von KI zur automatisierten Textbeurteilung exploriert und optimiert. Ein wesentliches Ziel dabei ist, Schüler*innen durch eine akkurate und detaillierte automatisierte Leistungsdiagnostik unmittelbar Feedback geben zu können und sie so beim Erwerb schriftsprachlicher Kompetenzen besser zu unterstützen. Die Implementierung und Evaluierung solcher Ansätze steht allerdings noch am Anfang. Die Beiträge der zweiten Hälfte des Symposiums richten den Fokus auf die Umsetzung von Ansätzen automatisierter Textbeurteilung und automatisierten Feedbacks.

Der erste Symposiumsbeitrag vergleicht die Urteilsgenauigkeit von Grundschullehrkräften und Grundschullehramtsstudierenden bei L2 Englischtexten Schweizer Schüler*innen anhand der Rang-, Niveau- und Differenzierungskomponente. Dabei wird ein analytisches Bewertungsschema verwendet. Der zweite Beitrag untersucht die Urteilsgenauigkeit von Lehrkräften und Lehramtsstudierenden anhand von L1 Texten von Schüler*innen der neunten Jahrgangsstufe. Zentraler Gegenstand dieser Untersuchung ist der Referenzrahmen als verzerrendes Urteilsmerkmal. Untersucht wird, inwieweit die Qualität von Vergleichstexten die Textbeurteilung beeinflusst. Der dritte Beitrag beschäftigt sich mit analytischer Textbeurteilung durch KI. Auch beim Trainieren von Algorithmen zur automatisierten Textbewertung stellen menschliche Expert*innenurteile die Grundlage dar, anhand derer die Maschine lernt. Die Güte der menschlichen Urteile bildet somit sowohl den Richtwert als auch die Grenze der angestrebten und erreichbaren Genauigkeit maschineller Bewertung. Im vierten Beitrag wird eine Schreibaufgabe vorgestellt, bei der automatisierte analytische Textbewertungen mit zufriedenstellender Güte gelungen sind. Diese

automatisierten Bewertungen dienten in der dargestellten Untersuchung als Grundlage für automatisiertes Feedback an Schüler*innen, dessen Wirkung in einem randomisierten Kontrollgruppendesign untersucht wird.

Die Beiträge werden in Bezug auf die Urteilsgenauigkeit bei Lehrpersonen, Expert*innen und KI diskutiert.

Beiträge des Symposiums

Herausforderung Textbeurteilung: Eine Analyse von Urteils(un)genauigkeit bei der Beurteilung englischer Schülertexte der Primarstufe

Ruth Trüb[1], Stefan Keller[2], Jens Möller[3], Julian Lohmann[3], Thorben Jansen[4]
[1]Pädagogische Hochschule FHNW, Schweiz, [2]Pädagogische Hochschule Zürich, Schweiz, [3]Christian-Albrechts-Universität zu Kiel, Deutschland, [4]Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel

Theoretischer Hintergrund

Diverse Studien zeigen, dass die Bewertung von Textqualität eine sehr anspruchsvolle Aufgabe ist und die Ausprägung gewisser Textmerkmale die Beurteilung erschwert oder sogar die Urteilsgenauigkeit anderer Merkmale beeinflusst (z. B. Rezaei & Lovorn, 2010; Wolfe et al., 2016). Im Kontext des Lernens von Englisch als Fremdsprache (EFL) auf der Gymnasialstufe haben Vögelin et al. (2018, 2019, 2020) gezeigt, dass die Qualität von Wortschatz, Orthografie und Textorganisation zu einem Halo-Effekt führen kann, dass also die Urteilsgenauigkeit anderer Textmerkmale durch sie beeinflusst wird. Nach unserer Erkenntnis gibt es bisher jedoch keine Studien, welche die Urteilsgenauigkeit bei Textbewertungen im EFL-Kontext der Primarstufe untersucht haben. Dass dies notwendig ist, zeigt eine Studie von Trüb (2022), welche die Schreibkompetenz junger EFL-Lernender untersuchte, und bei der während eines Standard-Settings ähnliche Unregelmäßigkeiten beobachtet wurden (S. 130ff). Die vorliegende Studie untersucht daher, wie die Urteils(un)genauigkeit bei der Bewertung von Texten dieser Zielstufe beschaffen ist, um daraus entsprechende Schlüsse für die Aus- und Weiterbildung von Lehrpersonen ziehen zu können.

Fragestellung

• Wie akkurat werden Texte junger EFL-Lernender der Primarstufe beurteilt und welche Unterschiede gibt es dabei in Bezug auf die unterschiedlichen Beurteilungskriterien?

Methode

Um diese Forschungsfragen zu beantworten, wurde eine experimentelle Studie durchgeführt, in der 147 Lehramtsstudierende und Lehrpersonen der Primarstufe jeweils vier randomisiert zugeteilte Texte von EFL-Lernenden der Primarstufe anhand von neun Skalen beurteilten (kommunikativer Effekt, Detailliertheitsgrad, Kohärenz, Kohäsion, Komplexität und Korrektheit von Syntax und Grammatik, Wortschatz, Orthografie und Zeichensetzung). Ihre Urteile wurden mit denjenigen von trainierten Expertinnen und Experten verglichen, um Aussagen über die Urteilsgenauigkeit machen zu können. Für die Analyse wurde das globale Abweichungsmaß sowie die Niveau-, Differenzierungs- und Rangkomponente der Urteilsgenauigkeit berechnet.

Ergebnisse und ihre Bedeutung

Das globale Abweichungsmaß zeigt eine durchschnittliche absolute Abweichung der Probandenurteile zwischen 0.56 und 0.98 für die unterschiedlichen Beurteilungskriterien. Diese Werte unterscheiden sich alle hoch signifikant und mit hohen Effektstärken vom Wert 0, welcher eine perfekte Übereinstimmung von Probanden- und Expertenurteilen repräsentiert. Demzufolge scheint bei allen Beurteilungskriterien ein Trainingsbedarf zu bestehen. Die Analyse der Niveaukomponente zeigt, dass insbesondere die Kohäsion und die Komplexität von Syntax und Grammatik von den Probandinnen und Probanden systematisch unterschätzt wird. Bei der Differenzierungskomponente zeigt sich, dass beim Kriterium «kommunikativer Effekt» die Breite der Skala durch die Probandinnen und Probanden deutlich stärker genutzt wird als durch die Expertinnen und Experten. Sie tendieren also dazu, extremere Urteile zu fällen als die Expertinnen und Experten. Bei der Rangkomponente weisen relativ tiefe Korrelationen zwischen 0.33 und 0.40 bei den Kriterien «kommunikativer Effekt», Detailliertheitsgrad, Kohärenz und Kohäsion darauf hin, dass es für die Probandinnen und Probanden schwierig ist, die Texte in eine ähnliche Rangordnung zu bringen wie die Expertinnen und Experten, also verlässlich zwischen Texten mit hoher und niedriger Qualität zu unterscheiden.

Die Resultate deuten auf gewisse Unterschiede in der Art der Urteilsgenauigkeit hin, zeigen aber insgesamt, dass bei allen Beurteilungskriterien Trainingsbedarf besteht. Daher wird ein online Trainingstool entwickelt, mit dem Lehramtsstudierende und Lehrpersonen das Beurteilen von Texten junger EFL-Lernender üben können, und mit dem sie auf Urteilsungenauigkeiten aufmerksam gemacht werden können.

“How did she do in comparison to the rest of the class?” – reference frame effects on text assessment

Frederike Strahl[1], Thorben Jansen[2], Jörg Kilian[1], Lea Vanselow[1], Jens Möller[1]
[1]Christian-Albrechts-Universität zu Kiel, Deutschland, [2]Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel

Theoretischer Hintergrund

Schreibkompetenzen von Schülerinnen und Schülern gelten als eine entscheidende Voraussetzung für schulischen, akademischen und beruflichen Erfolg (Crossley & McNamara, 2016; MacArthur, Graham & Fitzgerald, 2016) und somit als eine der Schlüsselkompetenzen für die gesellschaftliche Teilhabe von Personen (Becker-Mrotzek, 2014). Schreibkompetenzen werden zu einem großen Teil im Schulkontext geprägt und erlernt (Graham, 2019). Vor diesem Hintergrund stehen sie im Zusammenhang mit den Lehrpersonen. Die genaue und korrekte Beurteilung von Leistungen im Schulkontext bildet eine wichtige Voraussetzung für guten Schreibunterricht und individuelle Fördermaßnahmen (Elliott, Lee & Tollefson, 2001; Helwig, Anderson & Tindal, 2001¸ Hennes, 2020). Allerdings mangelt es oftmals an übereinstimmenden Beurteilungen desselben Textes durch verschiedene Lehrkräfte (Birkel & Birkel, 2002; Cooksey, Freebody & Wyatt-Smith, 2007; Skar & Jølle, 2017). Ein Grund für die fehlende Übereinstimmung könnte die Nutzung der sozialen anstelle der kriterialen Norm für die Bewertungen sein (Ingenkamp & Lissmann, 2008; Cooksey et al., 2007). Bei der sozialen Norm gilt die Referenzgruppe als Bezugsrahmen bei Beurteilungen. Die individuelle Leistung wird mit der Klassenleistung verglichen und demnach bewertet. Referenzrahmeneffekte wurden bisher in Bezug auf das akademische Selbstkonzept (Marsh, 1987; Marsh, Trautwein, Lüdtke, Baumert & Köller, 2007; Marsh et al., 2008) sowie die Einschätzung der Leistungsfähigkeit von

Lernenden nachgewiesen (Baumert, Trautwein & Artelt, 2003; Ingenkamp, 1969). Die Befunde legen nahe, dass auch die Beurteilung von konkreten Schülerleistungen wie der Textbewertung vom Referenzrahmen beeinflusst werden kann, was die vorliegenden Studien untersuchen.

Fragestellung

1. Beeinflussen Referenzrahmeneffekte die Beurteilung von konkreten Schülerleistungen wie Texten?

2. Wird die Textbeurteilung durch den Referenzrahmen negativ beeinflusst?

Methode

In der ersten Studie beurteilten insgesamt N = 227 Lehrpersonen (54.41 % weiblich), davon n = 136 Lehramtsstudierende (Alter: M = 25.62, SD = 3.56) und n = 91 erfahrene Lehrkräfte (Alter: M = 37.42, SD = 11.99), zehn randomisiert gezogene Schülertexte aus der neunten Klasse mit dem Schülerinventar (Jansen, Keller, Machts & Möller, 2019; Kaiser, Möller, Helm & Kunter, 2015). Die Texte stammten von der Datenbank Lernertexte (FD-LEX, Becker-Mrotzek & Grabowski, 2018, https://fd-lex.uni-koeln.de/login/auth). Es handelte sich um die Textsorte Bericht. Untersucht wurde der Einfluss der mittleren Textqualität, die über die aggregierten Expertenurteile der zehn Texte operationalisiert wurde, auf die mittlere Beurteilung der Texte durch die Lehrpersonen. Die Teilnehmenden bewerteten die Texte anhand sechs dichotomer Beurteilungsskalen, die anschließend zu einem Urteil aggregiert wurden.

Die Datenerhebung zur zweiten Studie beginnt Ende Oktober 2022. Die Studie soll die Ergebnisse der ersten Studie experimentalpsychologisch absichern. Dazu werden Lehramtsstudierende nach dem Zufallsprinzip zwei Gruppen zugeteilt und bewerten anschließend vier Schülertexte anhand einer ganzheitlichen sowie einer analytischen Skala. Beide Gruppen erhalten jeweils einen Standardtext von mittlerer Qualität sowie drei Texte von schlechter bzw. guter Qualität zur Beurteilung. Die Texte sowie die Expertenurteile wurden vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) für das Projekt ASSET-G (Assessing Students‘ Texts – German) zur Verfügung gestellt. Bei der Textsorte handelt es sich ebenfalls um Berichte.

Ergebnisse und ihre Bedeutung

Die Mehrebenenanalyse der Daten von Studie 1 ergab einen signifikant positiven Effekt der Textqualität ($\beta = 1.00$; $p < .001$) auf die Textbeurteilung. Die Textqualität fungiert als guter Prädiktor für die Beurteilungen der Lehrkräfte. Außerdem zeigte sich ein negativer Einfluss der mittleren Textqualität ($\beta = -.20$; $p < .001$) auf die mittlere Textbeurteilung bei erfahrenen Lehrkräften und Lehramtsstudierenden. Eine höhere mittlere Textqualität führt zu schlechteren Beurteilungen als eine niedrigere mittlere Textqualität. Die Ergebnisse der vorliegenden Studie zeigen, dass Schülerinnen und Schüler in leistungsstarken Klassen bei Aufsätzen womöglich schlechter bewertet werden und somit im Nachteil gegenüber Lernenden in leistungsschwachen Klassen sind. Bei der zweiten Studie erwarten wir für den Standardtext in der Gruppe mit den schlechten Vergleichstexten eine schlechtere Beurteilung als in der Gruppe mit den guten Vergleichstexten.

Automatisierte analytische Textbeurteilung – Exploration der Potentiale am Beispiel von Schüler*innentexten in der ersten Fremdsprache Englisch

Julian Lohmann[1], Fynn-Ole Junge[1], Stefan Keller[2], Johanna Fleckenstein[3], Thorben Jansen[4], Nils-Jonathan Schaller[4], Jens Möller[1]
[1]Christian-Albrechts-Universität zu Kiel, Deutschland, [2]Pädagogische Hochschule Zürich, Schweiz, [3]Universität Hildesheim, [4]Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel

Theoretischer Hintergrund

Die Beurteilung von Schüler*innentexten ist eine arbeits- und zeitintensive Tätigkeit von Lehrkräften. Die automatisierte Bewertung von Schreibleistungen stellt daher eine vielversprechende Alternative dar, um Schüler*innen künftig häufiger und schneller Feedback zur Qualität ihrer Schreibleistungen zu geben (Yan et al., 2020). Dies ist vor allem deshalb relevant, da hohe schriftsprachliche Kompetenzen für das Erlangen von Bildungsabschlüssen und Erfolg in der modernen Arbeitswelt elementar sind (z.B. Keller, 2013; Kultusministerkonferenz, 2014a, 2014b).

In den vergangenen Jahren konnte wiederholt gezeigt werden, dass Schüler*innentexte mithilfe von Machine Learning (ML)-Algorithmen akkurat bewertet werden können (z.B. Ramesh & Sanampudi, 2021; Uto, 2021). Diese Befunde beziehen sich bisher weitestgehend auf holistische Urteile, also die Vergabe einer Gesamtnote je Text (Ke & Ng, 2019). Der Wert solcher holistischen Bewertungen für Diagnostik und differenziertes Schüler*innen-Feedback ist allerdings limitiert (Kumar & Boulanger, 2021). So wird bei holistischen Bewertungen nicht zwischen verschiedenen Aspekten von Schreibleistungen, wie sprachlichen Fähigkeiten, struktureller Gestaltung und inhaltlicher Stringenz unterschieden (Condon & Elliot, 2022). Für detaillierte Diagnostik und Feedback ist es jedoch erforderlich, ebensolche Teilaspekte von Schreibleistungen, auch analytische Kriterien genannt, explizit zu berücksichtigen. Die sogenannte automatisierte analytische Textbeurteilung, auch automatisiertes trait scoring (ATS), berücksichtigt eben diese Teilaspekte. Im Kontext der Automatisierung von Diagnostik und Feedback bei Schreibleistungen ist ATS daher ein zentraler nächster Schritt (z.B. Horbach et al., 2017; Ke & Ng, 2019).

Um die Potentiale von ATS zu explorieren, verwendeten wir in der vorliegenden Studie einen Korpus von argumentativen Englischtexten, geschrieben von Schüler*innen der Jahrgangsstufe 11 aus Deutschland und der Schweiz. Mehr als 2300 dieser Texte wurden durch geschulte Expert*innen des Data Processing and Research Center Hamburg (DPC) der International Association for the Evaluation of Educational Achievement auf Basis verschiedener analytischer Skalen bewertet.

Forschungsfragen

1. Inwieweit stimmen die Expert*innen beim Bewerten der englischen Schüler*innentexte hinsichtlich verschiedener analytischer Kriterien überein? Gibt es zwischen den Kriterien Unterschiede in der Interraterübereinstimmung?

2. Wie akkurat können unterschiedliche ML-Algorithmen, die auf Basis der Texte und der Expert*innenurteile trainiert wurden, verschiedene analytische Scores vorhersagen?

3. Da ML Algorithmen Expert*innenurteile selten perfekt vorhersagen, ist von einem Verlust an Information auszugehen. Deshalb untersuchen wir, inwieweit die Vorhersage von unterschiedlichen analytischen Kriterien tatsächlich einen Mehrwert an Information im Vergleich zu holistischen Scores bieten kann.

Methoden

Im ersten Schritt wurden zunächst Interraterkorrelationen berechnet und zwischen den übergeordneten Skalen Sprache, Struktur und Inhalt, sowie hierarchisch untergeordneten Skalen zu detaillierteren Teilaspekten verglichen.

Im zweiten Schritt wurden für die drei übergeordneten analytischen Skalen, deren Reliabilität alle im akzeptablen Bereich lagen, ML-Algorithmen trainiert. Als Basis dienten sowohl N-Gramme (z.B. Hirschle, 2022) als auch weitere textbeschreibende Features, ermittelt mit dem Feature-Extraction-Tool CTAP (Chen & Meurers, 2016). Die über 600 Variablen im resultierenden Datensatz dienten den ATS-Modellen als Prädiktoren. In einem Kreuzvalidierungsprozess wurde zur Ermittlung der Güte der Vorhersage die Korrelation zwischen vorhergesagten Werten und Expert*innenurteil herangezogen.

Um im dritten Schritt den Verlust an Within-Text-Varianz durch das automatisierte Scoring abzuschätzen, wurde dieser Varianzparameter mit den Expert*innenurteilen und mit den ML-Scores berechnet und verglichen.

Ergebnisse

(1) die Interraterübereinstimmung zwischen den analytischen Skalen unterschieden sich signifikant. Außerdem schienen (2) verschiedene ML-Algorithmen unterschiedlich gut geeignet, die analytischen Bewertungen der Texte vorherzusagen. Darüber hinaus zeigten sich Unterschiede in der Vorhersagbarkeit zwischen den analytischen Skalen, die auch mit dem Ausmaß der Interraterübereinstimmung korrespondierten. Zuletzt zeigte sich (3) bei den drei übergeordneten analytischen Skalen kein bedeutsamer Verlust an Within-Text-Varianz.

Diskussion

Die Ergebnisse implizieren, dass mit gängigen ML-Algorithmen ein differenziertes Scoring von analytischen Beurteilungskriterien mit akzeptabler Genauigkeit und ohne nennenswerten Verlust von Within-Text-Varianz möglich ist. Dies setzt allerdings menschliche Beurteilungen mit ausreichender Interraterübereinstimmung voraus, da die Qualität und Genauigkeit der automatischen Scores bei diesem Ansatz maßgeblich von der Güte der Expert*innenurteile abhängt.

Schreibleistung maschinell beurteilen und fördern: Eine experimentelle Studie zur Wirksamkeit KI-basierten Feedbacks

Johanna Fleckenstein[1], Lucas Liebenow[2], Andrea Horbach[3], Jennifer Meyer[2]
[1]Universität Hildesheim, [2]Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel, [3]FernUniversität in Hagen

Theoretischer Hintergrund

Künstliche Intelligenz (KI) kann komplexe sprachproduktive Leistungen effizient beurteilen und damit eine Grundlage für die adaptive Förderung von Schreibkompetenzen schaffen. Basierend auf computerlinguistischen Algorithmen können didaktische Entscheidungen getroffen werden, die als individuelles Feedback den Lernenden im Schreibprozess zur Verfügung stehen.

In Lehr-Lern-Kontexten bezieht sich Feedback im Allgemeinen auf alle Informationen, die einer Person während oder nach einem Lernprozess gegeben werden. Feedback wirkt sich nur dann positiv auf die Leistung der Lernenden aus, wenn es Unsicherheit und kognitive Belastung verringert, indem es Informationen präsentiert, die zur Verbesserung der Aufgabenleistung erforderlich sind (Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Shute, 2008). Nach Narciss (2018) sollte

informatives Lernfeedback sowohl evaluative Informationen (d. h. Informationen über die aktuelle Aufgabenleistung) als auch tutorielle Informationen (d. h. ausführliche Informationen zur Verbesserung der Aufgabenleistung) enthalten, um das Lernen effektiv zu unterstützen.

Die Evidenz zur Wirksamkeit von automatischem, KI-basierten Feedback auf die Schreibleistung wird als heterogen beschrieben (McNamara et al., 2015; Stevenson & Phakiti, 2014; Strobl et al., 2019). Fleckenstein et al. (in press) zeigten in einem systematischen Überblick, dass die Effekte im Allgemeinen vielversprechend sind, aber es nur wenige Studien mit randomisiert-kontrollierten experimentellen Designs gibt und in diesen der Effekt individuellen Feedbacks häufig mit nicht-adaptiven Schreibförderungsmaßnahmen konfundiert ist.

Fragestellung

Dieser Beitrag stellt eine experimentelle Untersuchung der Effektivität informativen tutoriellen Feedbacks im Kontext automatischer Textbeurteilung dar. Als Voraussetzung wird die Qualität der maschinellen Urteile geprüft, indem diese mit menschlichen Urteilen verglichen werden. Anschließend wird die Effektivität des automatischen Feedbacks untersucht, indem die Leistungsentwicklung in einer digitalen Lernumgebung mit einer Kontrollgruppe verglichen wird.

Methode

Das Schreiben formaler E-Mails in englischer Sprache ist eine wichtige Kompetenz in gesellschaftlichen und beruflichen Kontexten und daher auch Thema im Englischunterricht der Sekundarstufe 1. Daher wurde diese Textsorte für die vorliegende Studie gewählt und zunächst ein Korpus mit 1104 E-Mails zu drei verschiedenen Aufgaben erhoben. Der Korpus wurde von zwei Expert:innen auf fünf Kriterien annotiert und anschließend unter Verwendung eines Gradient Boosting Classifiers maschinell kodiert. Der resultierende Algorithmus stellte die Basis für die experimentelle Feedback-Studie dar.

Die Wirksamkeit des auf den maschinellen Urteilen basierenden Feedbacks wurde in einem randomisiert-kontrollierten Experiment mit N = 105 Schüler:innen der Sekundarstufe 1 (50.5% weiblich; Alter: M = 14.41, SD = 0.81) untersucht. Die Lernenden bekamen die Aufgabe, in einer E-Mail an eine Sprachschule um zusätzliche Informationen zu bitten. Die E-Mails wurden anhand fünf Kriterien bewertet. Zu jedem Kriterium bekam die Experimentalgruppe ein Feedback, das sowohl evaluative (erfüllt/nicht erfüllt) als auch tutorielle (Tipps und Formulierungshilfen) Informationen enthielt. Der Kontrollgruppe wurden die Kriterien ohne weitere Informationen genannt. Die Schüler:innen wurden aufgefordert, ihren Text anhand der Kriterien bzw. des Feedbacks zu überarbeiten.

Ergebnisse und ihre Bedeutung

Die Analysen der maschinellen Urteile zeigten, dass eine automatische Bewertung der E-Mails hinreichend zuverlässig ist. Die Übereinstimmung mit den menschlichen Urteilen lag je nach Kriterium und Aufgabe zwischen 81 und 94 Prozent (QWK .35-.87). Im Vergleich dazu lag die Übereinstimmung zwischen den beiden menschlichen Urteilen zwischen 85 und 97 Prozent (QWK .43-.91).

Es wurde eine Mixed ANOVA durchgeführt, um die Leistungsverbesserung vom ersten Entwurf zur finalen Revision in der Experimental- und Kontrollbedingung zu vergleichen. Es zeigte sich ein deutlicher Effekt zugunsten der Experimentalgruppe ($F(1, 103)=11.92$, $p < .01$, $d = 0.68$). Eine weiterführende Analyse auf Ebene der Kriterien zeigte, dass die Leistungssteigerung insbesondere auf eine erfolgreiche Umsetzung bestimmter Kriterien zurückzuführen war.

Der vorliegende Beitrag zeigt, dass die KI-basierte Schreibförderung erfolgreich umgesetzt werden kann, sowohl hinsichtlich der Akkuratheit automatischer Urteile als auch der Wirksamkeit des darauf basierenden Feedbacks. Die Potenziale und Herausforderungen des Einsatzes automatischen Feedbacks zur Beurteilung und Förderung von Schreibleistungen im schulischen Kontext werden diskutiert.

H-17: Die Bedeutung der beruflichen Orientierung für einen erfolgreichen Übergang von der Schule in die Ausbildung

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* T03 R04 D10

Die Bedeutung der beruflichen Orientierung für einen erfolgreichen Übergang von der Schule in die Ausbildung

Chair(s): **Alexandra Wicht** (Uni Siegen/ BIBB), **Mona Granato** (BIBB, Germany), **Julia Gorges** (Philipps-Universität Marburg)
*Diskutant*in(nen):* **Birgit Ziegler** (TU Darmstadt Institut für Allgemeine Pädagogik und Berufspädagogik)

Der Ausbildungsstellenmarkt ist seit Jahren von zunehmenden Passungsproblemen geprägt. Neben regionalem und beruflichem Mismatch wurden qualifikatorische Mismatches als weitere Ursache für Disparitäten identifiziert. Disparitäten haben erhebliche Konsequenzen für den Einzelnen und die Ausbildungsbetriebe. Zum einen existieren Unterschiede in den Lebensläufen von Jugendlichen, die unmittelbar nach der Schule eine Ausbildung beginnen und solchen, die erst nach längerer Suche einen Ausbildungsplatz finden. Zum anderen geht die steigende Zahl unbesetzter Ausbildungsplätze mit einem Ressourcenverlust seitens der Unternehmen einher, der sich auch in bedeutenden Schwierigkeiten bei der Fachkräftesicherung niederschlägt.

Die beruflichen Orientierungen der Jugendlichen, d. h. ihre beruflichen Aspirationen, Interessen und Entscheidungen, sind ein *zentraler* Schlüssel zum Verständnis der Disparitäten auf dem Ausbildungsstellenmarkt. Aus einer interdisziplinären Perspektive geht das Symposium daher der Frage nach, welche Rolle berufliche Orientierungen für Übergangsgeschehen und -erfolg spielen und welche individuellen, sozial vermittelten, institutionellen und strukturellen Faktoren zu einer gelingenden Berufsorientierung und Ausbildungszugang beitragen.

Die Social Cognitive Career Theorie (SCCT) sieht den Zugang in Ausbildung, den Ausbildungserfolg und die berufliche Platzierung Jugendlicher als Prozess, der einerseits durch individuelle sowie sozial vermittelte Unterschiede in den beruflichen Orientierungen von Jugendlichen geprägt ist und andererseits durch unmittelbare Gelegenheitsstrukturen, die den Zugang zu Ausbildung und den Ausbildungserfolg fördern oder erschweren, darunter Unterstützungen und Hindernisse bei der Berufswahl und Charakteristika regional heterogener Ausbildungsmärkte. Individuelle (z.B. individuelle Fähigkeiten), sozial vermittelte Faktoren wie Gelegenheitsstrukturen und ihre individuelle Wahrnehmung können bei Jugendlichen in der Berufsorientierungsphase zu einer Einschränkung von Berufswahloptionen z.B. in geschlechtsspezifischer oder räumlicher Sicht beitragen.

Die ersten beiden Symposiumsbeiträge nehmen ausgewählte Faktoren in den Blick, die eine gelingende Berufsorientierung ausbildungssuchender Jugendlicher unterstützen oder erschweren. Dabei fokussiert der erste Beitrag auf die Bedeutung der Selbstwirksamkeit im Prozess der Berufsorientierung und untersucht, inwiefern Rollenvorbilder, die in Berufsorientierungsangebote eingebunden sind, dazu beitragen können die Selbstwirksamkeit von Schülern und insbesondere Schülerinnen im Hinblick auf die Wahl eines MINT-Berufes zu erhöhen, um damit der Tendenz entgegenzuwirken, dass Frauen seltener als Männer eine Laufbahn im MINT-Bereich aspirieren. Der zweite Beitrag nutzt pädagogisch-psychologische Zugänge in einem Mixed-Methods Design, um die Rolle von Lesekompetenz und Lesemotivation für die Berufsorientierung Jugendlicher zu beleuchten, die im ersten Anlauf keine Ausbildungsstelle erreicht haben und sich daher im Übergangssektor befinden. Der dritte Beitrag widmet sich aus soziologischer Perspektive der Frage nach der

Bedeutung des Zusammenspiels beruflicher Orientierungen, der regionalen Verfügbarkeit des angestrebten Berufs und sozial bindender Faktoren an den Heimatort für die Entscheidung Jugendlicher, eine Ausbildung außerhalb des Wohnortes beim Übergang in die Berufsausbildung aufzunehmen.

Beiträge des Symposiums

Warum junge Frauen seltener Fachinformatikerinnen werden möchten und wie Rollenvorbilder dies verändern könnten

Mona Granato, Alba Estela Esteve, Janina Beckmann
BIBB

Fragestellung

Trotz einer großen Vielfalt an Berufen konzentrieren sich die Berufsaspirationen Jugendlicher auf relativ wenige Berufe, was den Mismatch auf dem Ausbildungsstellenmarkt vertieft. Zudem ist die Berufswahl von Mädchen und Jungen weiterhin stark nach Geschlecht strukturiert, gerade im Pflegebereich oder bei (informations)technischen MINT-Berufen. Die horizontale Geschlechtersegregation erweist sich als ein Strukturmerkmal von Arbeits- und Ausbildungsmärkten: Der Frauenanteil in MINT-Ausbildungsberufen ist im vergangenen Jahrzehnt lediglich um 0,9 Prozentpunkte gestiegen und lag 2020 bei 11%. Besonders deutlich sind Geschlechterunterschiede im Ausbildungsberuf Fachinformatik – mit nur 9 % Frauen als Ausbildungsanfänger*innen (2021). Da MINT-Berufe mit günstigeren Beschäftigungs- und Aufstiegsaussichten verbunden sind, trägt eine geschlechtstypische Berufswahl zur Verfestigung vertikaler Geschlechterunterschiede bei. Entscheidend für gelingende Übergänge ist es daher Geschlechterunterschiede bei der Berufswahl im MINT-Bereich zu erklären und zu verringern. Dieser Beitrag untersucht, (1) warum junge Frauen seltener den Beruf Fachinformatiker*in anstreben und (2) wie Rollenvorbilder die geschlechtsspezifische Wahrnehmung des Berufs Fachinformatikerin verändern könnten.

Theorie

Bislang ist noch relativ unbekannt, welche Faktoren Geschlechterunterschiede in Bezug auf den Beruf Fachinformatiker*in erklären können. In diesem Beitrag liegt der Fokus dabei, in Anlehnung an die SCCT-Theorie, auf individuellen und sozial vermittelten Faktoren der Berufswahl.

In einem ersten Schritt beleuchten wir vier Erklärungsfaktoren: (1) Frauen könnten eine geringere berufsspezifische Selbstwirksamkeit aufweisen, also ein geringeres Vertrauen darin, die Anforderungen des Berufs zu bewältigen (Bandura, 1977). So unterschätzen Frauen ihre Fähigkeiten in Mathematik systematisch. (2) Frauen könnten sich weniger für die Tätigkeiten des Berufs interessieren, ein Erklärungsfaktor der auf empirische Befunde zu geschlechtsspezifischen Berufsinteressen rekuriert. (3) Geschlechteruntypische Berufe sind mit einer geringeren sozialen Anerkennung verbunden. Frauen könnten den Beruf Fachinformatiker*in aufgrund unzureichender sozialer Unterstützung durch Familie und Freunde ausschließen. (4) Zudem könnten, bedingt durch geschlechtsspezifische Sozialisationserfahrungen, die Kenntnisse über einen männlich-dominierten Beruf dahingehend geschlechtsspezifisch geprägt sein, dass Frauen über den Beruf Fachinformatiker*in weniger gut informiert sind.

In einem zweiten Schritt wird untersucht, wie Rollenvorbilder diese geschlechtsspezifische Wahrnehmung von Berufen verändern könnten. Die sozialkognitive Lerntheorie Bandura‘s (1977), als wesentlicher Bestandteil der SCCT-Theorie, geht davon aus, dass Rollenmodelle wesentlich über die bedeutsame Funktion des „vicarious learning“, d.h. durch das Lernen von Vorbildern, auf die Selbstwirksamkeit wirken. Berufliche Vorbilder könnten demnach dazu beitragen, die berufsbezogene Selbstwirksamkeit junger Frauen im Beruf „Fachinformatiker“ zu stärken. Des Weiteren wird untersucht, inwieweit berufliche Rollenvorbilder das Interesse an den Tätigkeiten, die wahrgenommene soziale Anerkennung und die Kenntnisse über den Beruf erhöhen können. Im Vordergrund steht dabei die Frage, ob (weibliche oder männliche) Rollenvorbilder einen stärkeren Effekt erzielen. Weibliche Rollenvorbilder könnten durch das Aufbrechen von Geschlechterstereotypen Schülerinnen ein größeres Identifikationspotential bieten.

Methode

Die Fragestellungen werden anhand der Daten der „BIBB-TUDa-Berufsorientierungsstudie“ analysiert. In diesem Projekt wird das Berufsorientierungsangebot „Ausbildungsbotschafter“ im Rahmen eines quasi-experimentellen, Längsschnittdesigns untersucht. Auszubildende stellen Schüler*innen auf „Augenhöhe“ ihren Ausbildungsberuf und -alltag vor und bieten ihnen als berufliche Rollenvorbilder Identifikationspotentiale – neben der Wissensvermittlung. Dieser Beitrag stützt sich auf eine Vorher-Nachher-Befragung von Schüler*innen, die Fragen zum Ausbildungsberuf Fachinformatiker*in beantwortet haben. Geschlechterunterschiede in der Wahrnehmung des Berufs werden mittels multivariater Analysen auf Basis der ersten Erhebungswelle geprüft (Regressionsverfahren, Dekompositionsanalysen). Ergebnisse zur Wirksamkeit von Rollenvorbildern im Beruf Fachinformatiker*in werden mittels Mittelwertvergleichen und Regressionsverfahren auf Basis der Vorher-Nachher-Befragung untersucht.

Ergebnisse

Erste Ergebnisse zeigen statistisch signifikante Unterschiede in der berufsbezogenen Selbstwirksamkeit, den Kenntnissen über den Beruf und den vermuteten Reaktionen des sozialen Umfelds beim Beruf Fachinformatiker*in. Analysen zur Wirksamkeit der Rollenvorbilder der Vorher-Nachher-Befragung weisen darauf hin, dass Rollenvorbilder die berufliche Selbstwirksamkeit im Beruf "Fachinformatiker*in" für beide Geschlechter erhöhen. Die folgenden Auswertungen untersuchen, inwieweit sich diese Effekte nach Geschlecht des Rollenvorbildes unterscheiden. Berufsorientierungsangebote mit Ausbildungsbotschafter*innen als Rollenvorbilder könnten sich über die Erhöhung der beruflichen Selbstwirksamkeit als erfolgsversprechender Weg erweisen, bestehende Engführungen auf wenige geschlechtstypische Berufe aufzubrechen und das Interesse junger Frauen an MINT-Berufen zu erhöhen.

Lesend zum passenden Beruf? Eine Mixed-Methods-Studie in der Berufsfachschule zum Übergang in Ausbildung (BÜA)

Jule Hangen[1], **Eveline Wuttke**[1], **Julia Gorges**[2]
[1]Goethe-Universität Frankfurt, [2]Philipps-Universität Marburg

Fragestellung, Theorie

Die 2017 in Hessen eingeführte Berufsfachschule zum Übergang in Ausbildung (BÜA) stellt die Berufsorientierung in den Mittelpunkt, um Schüler*innen den Übergang in ein Ausbildungsverhältnis zu ermöglichen. Betrachtet man die bisherige Schullaufbahn dieser Schüler*innen, kann als ein gemeinsames Hindernis beim Eintritt in die Berufsausbildung ihre geringe Lesekompetenz und Lesemotivation identifiziert werden. Aus Perspektive der Social Cognitive Career Theory kann ein geringes Leseniveau sowohl das Informieren über geeignete Ausbildungsberufe i.S. einer Erschließung der Gelegenheitsstrukturen als auch das erfolgreiche Durchlaufen eines Bewerbungsverfahren i.S. der Selbstwirksamkeit – entwickelt auf Basis bisheriger (schulischer) Erfolge – beeinträchtigen. Mit der vorliegenden Studie soll die Rolle des Lesens für die Berufsorientierung von Jugendlichen in BÜA näher beleuchtet und hierfür folgende Fragen adressiert werden:

1. In welchem Verhältnis stehen Lesemotivation, die Auswahl von Lesestoff und die Berufsorientierung?

2. Zeigt sich ein Zusammenhang zwischen selbstberichteter Lesekompetenz und Lesemotivation einerseits und dem Stand der Berufsorientierung der Schüler*innen andererseits?

Methode

Im Rahmen einer Mixed-Methode Studie bearbeiteten 87 Schüler*innen hessischer BÜA-Klassen (77,9% sprechen zu Hause ausschließlich oder auch Deutsch, 87,2% lesen täglich Texte in deutscher Sprache) einen standardisierten Fragebogen zu verschiedenen Aspekten des Lesens; 59 von ihnen nahmen im Anschluss zudem an einem Leitfadeninterview teil. Die Interviews wurden mittels qualitativer Inhaltsanalyse ausgewertet. Die quantitativen Daten wurden korrelativ sowie mittels t-Tests ausgewertet; Angaben zur aktuellen Berufsorientierung (kein konkreter Beruf oder weiterer Schulbesuch geplant, berufliche Domäne klar, Beruf klar) wurde hierfür auf Basis der inhaltsanalytischen Ergebnisse ergänzt.

Ergebnisse

Deskriptiv rangiert die Lesemotivation im mittleren Bereich einer fünfstufigen Antwortskala (Lesefreude: M=2.27; SD=.80; Leseinteresse: M=2.66, SD=0.84). Am häufigsten lesen die Schüler*innen persönliche Nachrichten (z.B. in Social-Media-Apps) und Internetseiten. Romane und Sachbücher werden häufiger gelesen als Zeitschriften und Tageszeitungen. Je stärker die Lesemotivation desto häufiger wird auch zum Vergnügen gelesen (Romane: rsp=.55; Sachbücher: rsp=.40; $p < .05$). Die in den Interviews am häufigsten genannten Themen, zu denen gelesen wird, sind aktuelle politische Themen und Informationen über die Berufsausbildung gefolgt von Biografien und Fantasy-Romanen.

Mehr als die Hälfte der Schüler*innen gaben an, dass sie das Internet nutzen, um sich über Berufe und Berufsausbildungen zu informieren. Weitere 55% informieren sich bei ihren Lehrkräften und etwa die Hälfte auch bei Freunden/ Bekannten und Eltern. Die Schüler*innen berichten jedoch, dass sie bei Internetrecherchen berufsspezifische Begriffe oft nicht verstünden, nicht wüssten, wie sie Suchmaschinen effektiv nutzen und die Vertrauenswürdigkeit von Websites beurteilen könnten, und

es für sie schwierig ist, einzuschätzen, ob sie Chancen auf eine Berufsausbildung im jeweiligen Beruf hätten.

In den Interviews gaben 29 Schüler*innen an, einen konkreten Berufswunsch zu haben, 12 gaben an, eine bestimmte Berufsrichtung zu präferieren, drei haben bereits einen Ausbildungsvertrag unterschrieben und 13 gaben an, weiter zur Schule gehen und einen höheren Abschluss machen zu wollen. Eine explorative Analyse der Zusammenhänge zwischen Berufsorientierung (konkreter Berufswunsch versus Domänenorientierung) und selbstberichteter Lesekompetenz (Fähigkeitsselbstkonzept, Deutschnote) sowie Lesemotivation (Lesefreude, Lernzielorientierung) zeigte, dass bessere Lesekompetenz und höhere Lesemotivation mit konkreteren Berufsvorstellungen einhergehen. Dieser Zusammenhang fiel jedoch nicht signifikant aus. Demgegenüber berichteten Schüler*innen mit konkreten Berufswunsch eine signifikant bessere Deutschnote ($t(35)=2.15$; $p<.05$). Allerdings waren die stärksten Leser*innen hinsichtlich Kompetenz und Motivation in der Gruppe zu finden, die einen höheren Schulabschluss anstrebte.

Diskussion

Ungeachtet ihrer bisherigen schulischen Laufbahn weisen die untersuchten BÜA-Schüler*innen eine mittlere Lesemotivation auf, nicht wenige lesen gerne und auch zum Vergnügen. Das Lesen von Texten, für das Schüler*innen vielfältige Quellen und Medien nutzen, leistet einen wichtigen Beitrag zur Berufsorientierung und erweist sich als relevant für die Konkretisierung beruflicher Pläne, stellt jedoch auch eine Herausforderung dar. So scheint es den Schüler*innen schwerzufallen, auf Basis verfügbarer Informationen einzuschätzen, in welchen Berufen sie Chancen auf einen Ausbildungsplatz hätten. Eine Leseförderung z.B. im Berufsbildenden Unterricht könnte zu einem Gelingen des Übergangs in Ausbildung beitragen.

»Should I stay or should I go?« Das Zusammenspiel von Berufsorientierungen, Gelegenheitsstrukturen und räumlicher Mobilität beim Übergang Jugendlicher in die Berufsausbildung

Linda Hoffmann[1], **Alexandra Wicht**[2]
[1]BIBB, [2]Uni Siegen/BIBB

Fragestellung

Die räumliche Mobilität von Jugendlichen ist einerseits ein wichtiges Mittel zur Lösung von Passungsproblemen auf Ausbildungs- und Arbeitsmärkten, die starke regionale Disparitäten aufweisen (Matthes & Ulrich, 2018). Andererseits ist räumliche Mobilität mit höheren monetären Erträgen auf individueller Ebene verbunden und daher eine Quelle sozialer Ungleichheit (Lemistre & Moreau, 2009). Während Analysen aggregierter Daten eine beträchtliche räumliche Mobilität von Auszubildenden aufzeigen (Bogai et al., 2008; Jost et al., 2019), ist die Forschung zu ihren Determinanten rar. Daher analysiert diese Studie die Bedeutung individueller beruflicher Orientierungen, sozial bindender Faktoren und regionaler Gelegenheitsstrukturen für die räumliche Mobilität von Jugendlichen beim Übergang von der Schule in die Berufsausbildung.

Theorie

Nach der SCCT (Lent et al., 2002) sind individuelle Aspirationen und subjektiv wahrgenommene Erfolgserwartungen, die durch externe Gelegenheitsstrukturen bedingt sind, von zentraler Bedeutung

für die erfolgreiche Umsetzung beruflicher Orientierungen. Insbesondere für Jugendliche mit hohen beruflichen Aspirationen bzw. geringen Erfolgserwartungen, etwa durch eine längere erfolglose Ausbildungsplatzsuche oder eine ungünstige Ausbildungsmarktlage vor Ort, kann räumliche Mobilität einen erfolgreichen Übergang in die Ausbildung ermöglichen.

Vorherrschende Theorien zur Erklärung räumlicher Mobilität stammen aus dem neoklassischen Rational-Choice-Rahmen, der berufs- oder ausbildungsbezogene Mobilität als primär wirtschaftliche Kosten-Nutzen-Entscheidung betrachtet, jedoch nicht-monetären Anreizen bzw. Hindernissen weniger Beachtung schenkt (Cadwallader, 1989). In Erweiterung dieser Perspektive und in Anlehnung an die SCCT integrieren wir daher theoretische Ansätze, die Mobilität als Ressource bzw. Potential – auch Motility genannt (Kaufmann et al., 2004) – ansehen und die Bedeutung individueller Motivation sowie struktureller Opportunitäten betonen. Neben der Rolle individueller beruflicher Orientierungen legen Studien insbesondere die Bedeutung sozial bindender Faktoren und regionaler Gelegenheitsstrukturen nahe, die die Mobilitätsbereitschaft mindern bzw. fördern können (Haldimann et al., 2021; Windzio, 2008).

Methode

Es wurden repräsentative Längsschnittdaten des Nationalen Bildungspanels NEPS (Startkohorte 4, N=3.143) (Blossfeld & Roßbach, 2019; NEPS Network, 2021) verwendet, um die Prädiktoren der räumlichen Mobilität von Schüler*innen zu analysieren, die innerhalb von zwei Jahren nach Verlassen des allgemeinbildenden Schulsystems in eine Berufsausbildung eintreten. Zur Abbildung regionaler Gelegenheitsstrukturen wurden administrative Regionaldaten auf Ebene des Schulstandortes unter Verwendung der NUTS-3-Regionalabgrenzung (BA, 2022; BBSR, 2022) angespielt.

Es wurden multinomiale logistische Regressionsmodelle geschätzt, um die Rolle von beruflichen Orientierungen der Jugendlichen (Statusaspirationen; Suchdauer), sozial bindender Faktoren (Zufriedenheit mit Freunden, und Familie; Mitgliedschaft in Organisationen) und der regionalen Gelegenheitsstruktur (Pendlersaldo; Anteil des angestrebten Berufssegments) für die räumliche Mobilität zu untersuchen. In Anlehnung an das Konzept funktionaler Arbeitsmarktregionen (Kosfeld & Werner, 2012) wurde zwischen räumlicher Mobilität innerhalb und zwischen regionalen Arbeitsmärkten unterschieden.

Ergebnisse

Unsere Ergebnisse zeigen eine starke Prävalenz räumlicher Mobilität unter Jugendlichen beim Übergang von der Schule in die Berufsausbildung: 16% sind innerhalb und 23% zwischen regionalen Arbeitsmärkten mobil. Zudem zeigt sich, dass sowohl individuelle berufliche Orientierungen, die soziale Einbettung sowie insbesondere regionale Gelegenheitsstrukturen entscheidend für die räumliche Mobilität der Jugendlichen sind.

Höhere Statusaspirationen und eine längere Suchdauer erhöhen die Wahrscheinlichkeit Jugendlicher für die Aufnahme einer Ausbildung innerhalb bzw. zwischen regionalen Arbeitsmärkten mobil zu werden. Bezüglich der sozialen Einbettung zeigt sich, dass Jugendliche mit einer höheren Zufriedenheit mit Freunden ihre Ausbildung eher in der Herkunftsregion beginnen und weniger über Arbeitsmarktregionen hinweg mobil werden. Ähnliches konnte für die regionalen

Gelegenheitsstrukturen festgestellt werden: Jugendliche in günstigeren Regionen beginnen ihre Ausbildung eher vor Ort und werden seltener mobil. Sowohl ein höherer Anteil des angestrebten Berufssegments in der Region sowie ein höheres Pendlersaldo verringern die Mobilität innerhalb Arbeitsmarktregionen, letzteres auch zwischen diesen.

Die Ergebnisse geben Aufschluss über die Prävalenz räumlicher Mobilität beim Übergang in die Berufsausbildung und was sie bedingt. Insbesondere die mobilitätsfördernde Rolle beruflicher Orientierungen und ungünstiger räumlicher Gelegenheitsstrukturen deuten auf die zentrale Bedeutung räumlicher Mobilität für erfolgreiche Ausbildungsübergänge im Kontext zunehmender Passungsprobleme und regionaler Disparitäten auf dem Ausbildungsmarkt hin.

H-18: Förderbezogene Diagnostik im inklusiven Mathematikunterricht

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* R11 T00 D01

Förderbezogene Diagnostik im inklusiven Mathematikunterricht

Chair(s): **Michael Besser** (Leuphana Universität Lüneburg, Deutschland), **Poldi Kuhl** (Leuphana Universität Lüneburg, Deutschland)
*Diskutant*in(nen):* **Timo Leuders** (Pädagogische Hochschule Freiburg, Deutschland)

Inklusive Schule hat das Ziel, alle Lernenden unabhängig von ihren individuellen Lern- und Leistungsvoraussetzungen bestmöglich im Lernprozess zu unterstützen. So verstandenes inklusives Unterrichten (des Fachs Mathematik) wird dabei von Lehrkräften jedoch als sehr herausfordernd wahrgenommen (Jütte & Lüken, 2021). Es stellt sich entsprechend unmittelbar die Frage nach Möglichkeiten der Gestaltung guten inklusiven (Fach-)Unterrichts, welcher auf der Basis der Professionalität der Lehrkräfte eine größtmögliche „Passung“ einer heterogenen Schülerschaft einerseits sowie dem praktischen Unterrichtsgeschehen andererseits gewährt. Eine zentrale Option, eine derartige Passung herzustellen, bietet die Idee der gezielten Individualisierung von Lehr-Lern Prozessen auf Grundlage der Nutzung diagnostischer Informationen, wie sie insbesondere auch im Rahmen formativer Assessments generiert werden. Allerdings gibt es aktuell kaum empirisch gesichertes Wissen darüber, wie genau solch diagnostische Prozesse in inklusiven und damit heterogenen Lernumgebungen im Fach Mathematik erfolgreich umgesetzt werden können.

Dieses Desiderat aufgreifend werden in diesem Symposium Beiträge zusammengetragen, die sich auf Basis empirischer Befunde mit der Erfassung oder Nutzung diagnostischer Informationen, bspw. aus formativen Assessments, im inklusiven Mathematikunterricht auseinandersetzen. Alle Projekte werden derzeit im Rahmen der Förderlinie „Förderbezogene Diagnostik in der inklusiven Bildung“ durch das BMBF gefördert. Im ersten Beitrag wird eine basiskompetenzorientierte Diagnostik für die Erfassung mathematischer Kompetenzen aller Kinder und Jugendlichen einschließlich Schüler*innen mit sonderpädagogischen Förderbedarfs erprobt. Dabei wird der Frage nachgegangen, inwiefern dieses Diagnostikverfahren bei Schüler*innen mit dem sonderpädagogischen Förderbedarf (SPF) Lernen eingesetzt werden kann und wie die Leistungen der Schüler*innen mit SPF Lernen sich im Vergleich zu den Schüler*innen ohne SPF einordnen lassen. Der zweite Beitrag befasst sich mit dem Zusammenhang zwischen sprachlichen und mathematischen Fähigkeiten. Vorgestellt werden erste Ergebnisse aus der Entwicklungs- und Forschungsphase eines Projekts zur Entwicklung, Erprobung und Evaluierung eines Diagnostikinstruments für inklusive Schulen zur Förderung von Mathematik- und Lesekompetenz am Übergang von der Grundschule zur Sekundarstufe I. Im dritten Beitrag werden Ergebnisse aus Expert*inneninterviews mit Mathematiklehrkräften einer seit mehreren Jahren inklusiv arbeitenden Oberschule in Niedersachsen vorgestellt. Dabei werden die derzeitigen unterrichtlichen Praktiken im Hinblick auf Diagnostik und Rückmeldungen als Facetten von formativem Assessment in den Blick genommen um herauszufinden, vor welchen Herausforderungen Lehrkräfte derzeit im inklusiven Mathematikunterricht stehen. Der vierte Beitrag schließlich beschäftigt sich mit den Effekten von zwei verschiedenen Ansätzen formativen Assessments (planned for interaction und curriculum-embedded assessment) auf die Entwicklung der mathematischen Kompetenzen, das schulische Selbstkonzept und die kognitive Motivation von Grundschüler*innen.

In der abschließenden Diskussion werden die Beiträge aus Mathematikdidaktischer Sicht hinsichtlich möglicher Schlussfolgerungen für die förderbezogene Diagnostik eingeordnet.

Beiträge des Symposiums

Entwicklungsorientierte Diagnostik von Zahl-Größen-Kompetenzen bei Schüler*innen mit sonderpädagogischem Förderbedarf Lernen

Jana Christl, Ronja Steinhauser, Melanie Besca, Kristin Krajewski, Marco Ennemoser
Pädagogische Hochschule Ludwigsburg, Deutschland

Theoretischer Hintergrund: Mathematische Kompetenzen sind eine bedeutsame Schlüsselkompetenz für die Teilhabe am gesellschaftlichen Leben (Rammstedt, 2013; Sälzer et al., 2013). Um deren Entwicklung optimal fördern zu können, sind diagnostische Informationen über den individuellen Leistungsstand der Schüler*innen von zentraler Bedeutung. Im Sekundarstufenbereich stellt eine fundierte Diagnostik mathematischer Kompetenzen eine gewaltige Herausforderung dar (Moser Opitz, 2010), da von der 5. bis zur 13. Klassenstufe über verschiedene Schulformen und variierende Bildungspläne hinweg ein enormes Kompetenzspektrum abgedeckt werden muss, das in einem inklusiven Setting sogar noch deutlich breiter wird. Im hier vorgestellten Projekt wird daher eine basiskompetenzorientierte Diagnostik erprobt, die eine Erfassung mathematischer Kompetenzen bei Kindern und Jugendlichen aller Schularten einschließlich Schüler*innen mit sonderpädagogischem Förderbedarf (SPF) möglich machen soll. Vorherige Studien zeigten bereits, dass sich die erfassten mathematischen Basiskompetenzen an den allgemeinen Schulen über die Sekundarstufe hinweg systematisch weiterentwickeln und wesentlich mit curricularen mathematischen Kompetenzen assoziiert sind (Ennemoser et al., 2011; Krajewski et al., 2010). In der vorliegenden Arbeit wurde nun untersucht, ob dieses Diagnostikverfahren mathematischer Basiskompetenzen auch bei Schüler*innen mit SPF Lernen eingesetzt werden kann und inwiefern sich die Leistungen dieser Schüler*innen im Vergleich zu den Leistungen der Schüler*innen an allgemeinen Schulen einordnen lassen. Die vorliegende Untersuchung ist in die Längsschnittstudie „EnDiMath – Entwicklungsorientierte Diagnostik mathematischer Basiskompetenzen“ eingebettet, in welcher über einen Zeitraum von drei Jahren über alle Schularten hinweg die Entwicklung von Schüler*innen, welche anfänglich die 5. und 7. Klassen besuchten, in den mathematischen Basiskompetenzen begleitet wird.

Methode: Zum ersten Messzeitpunkt bearbeiteten N = 1161 Schüler*innen aus allgemeinen Schulen (darunter inklusiv beschulte Schüler*innen) sowie exklusiv beschulte Schüler*innen

mit SPF Lernen einen Test zur Erfassung mathematischer Basiskompetenzen, der die Zahl-Größen-Kompetenzen und die Fähigkeit zum schnellen Abruf einfacher Rechenfakten (Faktenabruf) abprüfte. Außerdem wurden fluide und kristalline Intelligenz über den CFT 20-R (Weiß, 2006) erhoben.

Ergebnisse: Erste Ergebnisse deuten darauf hin, dass das Diagnostikverfahren auch bei Schüler*innen mit SPF Lernen eingesetzt werden kann und neben einer Vergleichbarkeit mit gleichaltrigen Schüler*innen auch eine Differenzierung zwischen Schüler*innen mit SPF Lernen leisten kann. Wie erwartet, schneiden die Schüler*innen mit SPF Lernen hinsichtlich ihrer Zahl-Größen-Kompetenzen im Vergleich zu ihren Peers schwächer ab. Es zeigt sich jedoch ein positiver Entwicklungstrend von der 5. zur 7. Klasse. Dies weist darauf hin, dass die Schüler*innen mit SPF

Lernen, genau wie die Schüler*innen ohne SPF, über die Sekundarstufe hinweg eine positive Entwicklung hinsichtlich ihrer Zahl-Größen-Kompetenzen durchlaufen. Mit fluider sowie kristalliner Intelligenz und der Leistung im Bereich Zahl-Größen-Kompetenzen bestehen positive Zusammenhänge, die ähnlich ausgeprägt sind wie bei Schüler*innen allgemeiner Schulen. Chancen und Herausforderungen einer basiskompetenzorientierten Diagnostik bei Schüler*innen mit SPF Lernen im Fachbereich Mathematik werden diskutiert.

Förderbezogene Diagnostik in Mathematik inklusive Lesen - Erste Ergebnisse aus dem Projekt FORMEL

Stephan Hußmann[1], Anke Hußmann[1], Meike Mentrop[1], Marcus Nührenbörger[2], Lina Kortüm[2]
[1]Technische Universität Dortmund, Deutschland, [2]Westfälische Wilhelms-Universität Münster, Deutschland

Befunde aus zahlreichen Studien bestätigen seit Jahren den Zusammenhang von sprachlichen und mathematischen Fähigkeiten von Schüler*innen (Heinze et al., 2011; Mann Koepke & Miller, 2013; Martin & Multis, 2013; Schröder & Ritterfeld, 2014; Ufer et al., 2013; Walzebug, 2014, 2015). Für die Entwicklung des mathematischen Verständnisses sind sprachliche Kompetenzen der Schüler*innen elementar (Prediger et al., 2015) und gelten als Prädiktor für den Lernzuwachs (Paetsch et al., 2016). Unter den sprachlichen Fähigkeiten gilt das Leseverstehen als Schlüsselkompetenz (Grimm, 2008). Schwierigkeiten beim Lesen können zu Problemen bei der Erschließung von mathematischen Aufgabentexten und bei der Sinnkonstruktion führen (Prediger et al., 2019; Stephany, 2018). Es ist aber unklar, wie sich Lesekompetenzen auf die Entwicklung mathematischer Kompetenzen auswirken, welche Relevanz dabei hierarchieniedrige und -höhere Leseprozesse haben und welche Bedeutung der sozial-situativen Einbindung zukommt (Moser Opitz et al., 2018; Rosebrock & Nix, 2017; Thomas et al., 2015). Zugleich bestätigen Befunde aus IGLU, TIMSS und dem IQB-Bildungstrend den dringenden Bedarf an validen Diagnose- und Fördermaterialien für den Unterricht, da der Anteil an Kindern am Ende der Grundschulzeit mit nur rudimentären Lese- und Mathematikkompetenzen steigt (Hußmann et al., 2017; Selter et al., 2020; Stanat et al., 2022).

Das Projekt FORMEL setzt an diesem Bedarf an. Entwickelt, erprobt und evaluiert wird ein kombiniertes Diagnostikinstrument zur Förderung von Mathematik- und Lesekompetenz in inklusiven Schulen am Übergang von der Grundschule zur Sekundarstufe I.

Leitende Fragen sind u.a., wie sich sprachliche und mathematische Kompetenzen im mathematischen Lernprozess beeinflussen, wie sich ein förderbezogenes Diagnostikkonzept inklusive adaptiver Fördermöglichkeiten gestalten lässt, wie Lernprozesse der Schüler*innen verlaufen, welche Verstehensschwierigkeiten auftreten, und natürlich auch, wie Lehrer*innen die zu entwickelnde förderbezogene Diagnostik bestmöglich in ihrem Unterricht einsetzen können. Genutzt werden fachdidaktische, sonderpädagogische und pädagogisch-psychologische Perspektiven im Dialog mit Akteuren aus Praxis und Wissenschaft, um mathematische Kompetenzen und Lesefähigkeiten in ihren Einzelaspekten und in ihrem Zusammenspiel bearbeiten zu können.

Die förderbezogene Diagnostik, die am Ende als FORMEL-Konzept frei verfügbar sein soll, wird mathematische Förderung (Fokus auf kontextgestützte Aufgaben zum Operationsverständnis; Multiplikation und Division) mit Hilfestellungen aus der Leseforschung verknüpfen und gezielt individuelle Lese- und Mathematikkompetenzen der Schüler*innen berücksichtigen.

Das FORMEL-Konzept wird im Rahmen der designorientierten Entwicklungsforschung (Hußmann et al., 2013; Nührenbörger et al., 2019) in drei Phasen umgesetzt: der Entwicklung, der Forschung und der Dissemination des Diagnostikinstruments. Im Vortrag werden erste Ergebnisse aus der Entwicklungs- und der Forschungsphase vorgestellt. Seit Projektbeginn werden vor allem Diagnose-, aber auch Förderaufgaben mit Blick auf sprachliche Besonderheiten, schwierigkeitsgenerierende Merkmale, individuelle Rezeptionsprozesse, ebenso wie grafische und symbolische Darstellungen (weiter-)entwickelt und in Klassen- und Einzelsettings mit Schülerinnen in 4. und 5. Jahrgangsstufen an Schulen in NRW erprobt. Dazu werden in zyklisch aufeinander aufbauenden Designexperimenten den Schüler*innen die (weiter-)entwickelten Aufgabenformate im Rahmen von videographierten, leitfadengestützten diagnostischen Interviews vorgelegt.

Umsetzung von Leistungsdiagnostik und Leistungsrückmeldung im inklusiven Mathematikunterricht

Fynn Töllner, Kyra Renftel, Poldi Kuhl, Michael Besser
Leuphana Universität Lüneburg, Deutschland

Guter inklusiver (Mathematik-) Unterricht orientiert sich an den individuellen Lernvoraussetzungen der Schüler*innen. Dies begründet sich unter anderem dadurch, dass eine adaptive Kompetenz von Lehrkräften den Lernerfolg der Schüler*innen nachweisbar positiv beeinflussen kann (Brühwiler, 2014). Allerdings ist die Durchführung eines adaptiven Unterrichts mit einer individuellen Förderung aller Schüler*innen für Lehrkräfte eine Herausforderung (Stebler & Reusser, 2017). Eine vielversprechende Möglichkeit, den Unterricht adaptiv an die Bedürfnisse der Schüler*innen anzupassen und dieser Herausforderung in der Praxis zu begegnen, stellt formatives Assessment dar (Schütze et al., 2018). Dieses umfasst dabei insbesondere das kontinuierliche Diagnostizieren und Rückmelden von Schüler*innenleistungen, welches den Lernprozess wie selbstverständlich begleitet (Heritage, 2007). Wie genau ein derart gestaltetes Diagnostizieren und Rückmelden im inklusiven (Mathematik-)Unterricht erfolgreich und lernförderlich umgesetzt werden kann, ist jedoch weitestgehend unklar und muss entsprechend als Forschungsdesiderat angesehen werden.

Ausgehend von diesen Überlegungen untersucht das Projekt LERN-IF Möglichkeiten zur Implementation von lernprozessbegleitender Diagnostik und lernförderlichen Leistungsrückmeldung im inklusiven Fachunterricht am Beispiel des Unterrichtsfaches Mathematik. Mit dem Ziel, aktuelle schulische Praxis zu beschreiben und zu verstehen, soll in einem ersten Schritt nachgezeichnet werden, wie Lehrkräfte im inklusiven Mathematikunterricht die Leistungen ihrer Schüler*innen erfassen und zurückmelden. Aufbauend auf den gewonnenen Erkenntnissen wird in einem nächsten Schritt ein Konzept formativen Assessments für den inklusiven (Mathematik-)Unterricht erarbeitet und wissenschaftlich erprobt. Der vorliegende Beitrag greift erstgenanntes Ziel auf und geht den folgenden Forschungsfragen nach: (1) Wie setzen Lehrkräfte individuelle Diagnose und Rückmeldung im inklusiven Mathematikunterricht explizit um? (2) Welche konkreten Herausforderungen beschreiben die Lehrkräfte dabei für ihren alltäglichen Mathematikunterricht?

Um diese Fragen beantworten zu können, wurden im Sommer 2022 sechs halbstrukturierte Expert*inneninterviews mit N = 6 erfahrenen Mathematiklehrkräften (4w, 2m, 0d) einer seit mehreren Jahren inklusiv arbeitenden Oberschule in Niedersachsen geführt. Die Interviews dauerten je ca. eine

Stunde und erfragten derzeitige Praktiken zur Leistungsdiagnostik und - rückmeldung sowie wahrgenommene Chancen und Herausforderungen dieser Praktiken für einen inklusiven Fachunterricht. Die Expert*inneninterviews wurden mithilfe einer qualitativen Inhaltsanalyse nach Mayring (2015) mittels deduktiver Kategorienbildung entlang der zentralen Forschungsfragen von zwei Kodierer*innen unabhängig voneinander ausgewertet.

Die Ergebnisse liefern zentrale Einblicke zur aktuellen schulischen Umsetzung formativen Assessments im inklusiven Mathematikunterricht. So kann konstatiert werden: (1) Ein von den Lehrkräften zielgerichtet entwickeltes und auf dem bewussten Einsatz von expliziten Materialien für inklusiven Fachunterricht basierendes Konzept lässt sich auch nach mehreren Jahren inklusiver Unterrichtserfahrung in der schulischen Praxis nicht finden/beschreiben. (2) Als Herausfordernd bei der Durchführung von Diagnostik und Rückmeldung im inklusiven Mathematikunterricht erleben Lehrkräfte beispielsweise den Mangel an persönlichen Ressourcen wie z.B. Zeit.

Die Ergebnisse belegen einmal mehr die zwingende Notwendigkeit, fundamentale Entwicklungen des Schulsystems (wie die der inklusiven Beschulung aller Lernenden) gezielt durch evidenzbasierte Konzepte zur Unterrichtsentwicklung sowie zur professionellen Entwicklung zu begleiten bzw. zu unterstützen. Dies aufgreifend werden Implikationen für die (fachdidaktische) Umsetzung formativen Assessment im alltäglichen inklusiven (Fach-)Unterricht vorgestellt und diskutiert. Leitend ist dabei die Frage, welche Faktoren für die Entwicklung und Erprobung von Konzepten zur Implementation formativen Assessments zentral erscheinen, damit diese sowohl die Bedürfnisse der Lehrkräfte berücksichtigen als auch individuelle Lernprozesse der Schüler*innen unterstützen können.

Effekte zweier Ansätze formativen Assessments auf die mathematischen Kompetenzen, das schulische Selbstkonzept und die kognitive Motivation von Grundschüler:innen

Larissa Aust[1], Luise Eichholz[2], Petra Heß[2], Jeanne-Celine Linker[2], Marcus Nührenbörger[1], Jana Schiffer[1], Christoph Selter[2], Franz Wember[2], Elmar Souvignier[1]
[1]Westfälische Wilhelms-Universität Münster, Deutschland, [2]Technische Universität Dortmund, Deutschland

Beim formativen Assessment werden – begleitend zum Lernprozess – diagnostische Informationen über die Lernstände und den Lernfortschritt von Kindern erhoben. Auf dieser Grundlage soll eine hohe Passung zwischen Lernstand und Förderung und damit eine Optimierung individuellen Lernens ermöglicht werden (Black und Wiliam, 1998; 2009). Positive Effekte formativen Assessments auf die Leistungsentwicklung von Schüler:innen wurden generell belegt (z.B. Black und Wiliam, 1998; Klute et al., 2017; Lane et al., 2019). Auch erste Untersuchungen zu Wirkungen formativen Assessments auf motivationale Variablen zeigten positive Effekte (Hondrich et al., 2018; Leenknecht et al., 2020). Metaanalysen weisen jedoch darauf hin, dass die Wirkung formativen Assessments u.a. vom jeweiligen Schulfach und der konkreten Nutzung der Assessment-Informationen abhängt (Kingston und Nash, 2011; Klute et al., 2017; Lane et al., 2019). Ein Kriterium, anhand dessen die konkrete Nutzung formativen Assessments klassifiziert werden kann, ist das Ausmaß an Formalisierung und Strukturierung des Assessments. Beim Ansatz planned-for-interaction assessment (PI) plant die Lehrkraft Unterrichtssequenzen im Vorhinein und setzt gezielte diagnostische Impulse, um den Lernstand der Lernenden zu ermitteln. Bei dem stärker formalisierten Ansatz curriculum-embedded assessment (CE) werden bestehende oder selbst entwickelte diagnostische Instrumente genutzt, um systematisch den Lernstand der Schüler:innen zu erfassen (Heritage, 2007; Shavelson et al., 2008).

Das Ziel der vorliegenden Studie liegt darin, eine vergleichende Evaluation dieser beiden Ansätze im Anfangsunterricht in Mathematik vorzunehmen.

Fragestellung

Welchen Effekt haben zwei unterschiedlich stark strukturierte Ansätze (CE vs. PI) formativen Assessments auf die Entwicklung der mathematischen Kompetenzen, des schulischen Selbstkonzepts und der kognitiven Motivation von Grundschüler:innen?

Methode

An den Erhebungen nahmen Schüler:innen aus 67 Klassen des zweiten Schulbesuchsjahres in Nordrhein-Westfalen teil. Die Studie fand in einem Zweigruppen-Prä-Post-Design statt. Die Lehrkräfte beider Gruppen besuchten jeweils sechs Fortbildungsveranstaltungen über ein Schuljahr hinweg, in denen ihnen das Prinzip diagnosebasierter differenzierter Förderung vermittelt wurde. Zwischen den Fortbildungen erprobten sie konkrete Praxisaufgaben zu den jeweiligen Konzepten (CE vs. PI) in ihren Klassen. Lehrkräfte in der Gruppe CE setzten hierbei vorbereitete Diagnoseaufgaben (s.g. Standortbestimmungen) zur Erfassung des aktuellen Lernstandes ein und wählten auf der Basis dieser Ergebnisse passende Förderaufgaben aus. Lehrkräfte der Gruppe PI lernten, mittels reichhaltiger Aufgaben Diagnosegespräche zu führen und darauf aufbauend Förderaktivitäten zu entwickeln. Die Effekte der beiden Ansätze wurden auf der Ebene der Schüler:innen untersucht. Dazu fanden Erhebungen zu Beginn und am Ende des Schuljahres statt. Die mathematischen Basiskompetenzen wurden mithilfe des Basis-Math 1+ (Moser-Opitz et al., i.V.) bzw. Basis-Math-G 2+ (Moser-Opitz et al., 2020) erfasst. Das schulische Selbstkonzept wurde durch eine gekürzte Version der entsprechenden Skala des FEESS 1-2 (Rauer und Schuck, 2004) und die kognitive Motivation mit dem NFC-KIDS (Preckel und Strobel, 2017) erhoben.

Ergebnisse und Implikationen

Bisher wurden die Daten einer Teilstichprobe von N = 518 Kindern unter Kontrolle von Prätest-Unterschieden varianzanalytisch ausgewertet (eine mehrebenenanalytische Auswertung unter Berücksichtigung der Umsetzungstreue erfolgt, wenn der gesamte Datensatz aufbereitet ist). Dabei zeigten sich signifikante Unterschiede mit kleinen Effekten zugunsten der Gruppe PI im Hinblick auf die Mathematikleistungen ($F(1,509) = 14.88$, $p < .001$, partielles $\eta^2 = .028$) und die kognitive Motivation der Kinder ($F(1,515) = 3.91$, $p < .05$, partielles $\eta^2 = .008$). Bezüglich des schulischen Selbstkonzeptes konnten keine Unterschiede gefunden werden ($F(1,502) = 1.00$, $p > .05$, partielles $\eta^2 = .028$).

Die Ergebnisse deuten einen leichten Vorteil einer etwas weniger stark formalisierten und stärker individualisierten Form der Diagnostik an. Durch den Vergleich der beiden unterschiedlichen Konzepte können wichtige Hinweise für die konkrete Ausgestaltung formativen Assessments, insbesondere der Diagnostikkomponente, im Mathematikunterricht der Grundschule unter Berücksichtigung verschiedener Zieldimensionen gewonnen werden.

H-19: Informelles Lernen von Lehrpersonen in Deutschland – Analoge und digitale Perspektiven der Professionalisierung

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* R11 T00 D03

Informelles Lernen von Lehrpersonen in Deutschland – Analoge und digitale Perspektiven der Professionalisierung

Chair(s): **Sebastian Röhl** (Eberhard Karls Universität Tübingen, Germany), **Colin Cramer** (Eberhard Karls Universität Tübingen, Germany)
*Diskutant*in(nen):* **Cornelia Gräsel** (Bergische Universität Wuppertal)

Für eine erfolgreiche, lebenslange Tätigkeit als Lehrkraft sind die in den qualifizierenden universitären und praxisbezogenen Ausbildungsphasen erworbenen Kompetenzen nicht ausreichend und ein lebenslanges berufsbezogenes Lernen notwendig (Richter & Richter, 2020). Hierbei nimmt neben dem Lernen in organisierten Fort- und Weiterbildungen auch das Lernen am Arbeitsplatz eine wichtige Rolle ein (Schmidt-Hertha, 2020). So wird davon ausgegangen, dass etwa 80% der berufsbezogenen Fertigkeiten und des Wissens durch Lernprozesse außerhalb der institutionellen Lernveranstaltungen erworben werden (Coffield, 2000), wobei hier zwischen zufälligem inzidentellem und intentionalem informellem Lernen unterschieden werden kann (Dobischat & Gnahs, 2008).

Dieses Symposium fokussiert die Nutzung von intentionalen informellen Lerngelegenheiten durch Lehrpersonen, zu dem beispielsweise die bewusste Reflexion von Unterricht, das Einholen von Feedback, der kollegiale Austausch aber auch das Rezipieren von unterrichtsbezogenen Medien sowie das Erproben neuer Methoden und Unterrichtsformen gezählt werden (Kyndt et al., 2016). Bisherige Befunde zum informellen Lernen von Lehrpersonen in Deutschland zielten nur auf einzelne Lernformen (z.B. Grosche et al., 2020; Richter et al., 2011). Weiterhin ermöglichen die zunehmenden Möglichkeiten des Internets zur Informationssuche und zum Austausch erweiterte Formen des informellen Lernens (z.B. Fütterer et al., 2021). Durch die Beiträge des Symposiums werden sowohl für Deutschland repräsentative Befunde zur Nutzung von informellen Lerngelegenheiten in der Breite als auch zu spezifischen Entwicklungen, etwa zur Nutzung verschiedener Internetplattformen, vorgestellt.

Im ersten Beitrag wird die Konstruktvalidierung einer deutschen Adaption eines Instruments zur Erhebung der Nutzung von informellen Lerngelegenheiten durch Lehrpersonen vorgestellt. Die Befunde weisen darauf hin, dass das berufsbezogene Nutzungsverhalten von Online-Medien nicht mit der traditionell üblichen Rezeption von Printmedien gemeinsam betrachtet werden kann. Vertiefende Analysen zu den Effekten von demografischen, dispositionellen und auf den Beruf bezogenen Merkmalen auf die Lerngelegenheitsnutzung weisen u.a. darauf hin, dass die berufsbezogene Nutzung von Online-Medien nicht vom Alter abhängig zu sein scheint.

Daran anschließend thematisiert der zweite Beitrag, inwiefern Lehrpersonen die Plattform Instagram nicht nur zum Suchen und Bereitstellen von Unterrichtsmaterialien, sondern auch zur Ko-Kreation neuer Materialen nutzen. Es zeigt sich, dass Instagram für Lehrpersonen durchaus ein Ort des Austauschs sein kann, der etwa das Erleben emotionaler Unterstützung ermöglicht, wenngleich dieser bisher nur in einem geringen Maße stattfindet.

Mit dem Austausch von Materialien innerhalb von Twitter-Lehrkräfte-Communities beschäftigt sich der dritte Beitrag, der hierzu etwa anderthalb Millionen Tweets hinsichtlich der Verfassenden und der zur Verfügung gestellten Materialien analysiert.

Im Anschluss werden die Beiträge im Kontext der Bedeutsamkeit für die evidenzorientierte Lehrer:innenbildung und für die Nutzung informeller Lerngelegenheiten notwendiger Kompetenzen diskutiert.

Beiträge des Symposiums

Nutzung und Bedingungsfaktoren informeller Lerngelegenheiten von Lehrpersonen in Deutschland

Sebastian Röhl[1], **Jana Groß Ophoff**[2], **Colin Cramer**[1]
[1]Eberhard Karls Universität Tübingen, Germany, [2]Pädagogische Hochschule Vorarlberg, Austria

Im Gegensatz zur Erfassung der Teilnahmehäufigkeiten von Fort- und Weiterbildungen erweist sich eine möglichst umfassende Erfassung der Nutzung informeller Lerngelegenheiten bisher in der Breite als schwierig, da viele Erhebungsinstrumente mehr oder weniger zufällig als lernförderlich geltende Aktivitäten auflisten (Dobischat & Gnahs, 2008). Für den Lehrer:innenberuf identifizierte Kwakman (2003) in einer qualitativen Interviewstudie eine Vielzahl informeller Lerngelegenheiten, die in einem mehrschrittigen Verfahren schließlich zur Entwicklung und Validierung des Erhebungsinstruments Teachers‘ Professional Development at Work (TPD) (Evers et al., 2016) mit den Dimensionen ‘Experimenting’, ‘Collaborating for School Development’, ‘Collaborating for Lessen Development’, ‘Keeping Up-to-date: Reading’ sowie ‘Reflecting and Asking for Feedback’ führte. Die Erprobung einer deutschen Adaption dieses Instruments bildet die Grundlage für die hier dargestellte Untersuchung, wobei darüber hinaus die Zusammenhänge zwischen den Lerngelegenheitsdimensionen und verschiedenen, exemplarisch ausgewählten demografischen, dispositionellen und auf den Beruf bezogenen Merkmalen untersucht werden, die eine hohe Relevanz für die Lerngelegenheitsnutzung aufweisen (u.a. Cerasoli et al., 2018). Folgende Forschungsfragen werden adressiert:

FF1: Bewähren sich die bei Evers et al. (2016) identifizierten Dimensionen der Nutzung informeller Lerngelegenheiten in einer deutschen Übersetzung des Instruments?

FF2: Wie häufig nutzen Lehrpersonen in Deutschland verschiedene Dimensionen informeller Lerngelegenheiten?

FF3: Welche Effekte zeigen Alter, Geschlecht, fehlende Qualifikation (Quer-/Seiteneinstieg), allgemeine Selbstwirksamkeitserwartung und die Übernahme einer Funktionsstelle unter Kontrolle der Schulform auf die Nutzungshäufigkeit informeller Lerngelegenheiten von Lehrpersonen?

Methode

Für diese Studie wurde im Herbst 2019 durch den Felddienstleister forsa GmbH eine repräsentative Stichprobe von N = 405 Lehrpersonen aus Deutschland befragt. Grundlage des Online-Fragebogens bildete eine deutsche Adaption des TPD (21 Items, Antwortskala 1 = nie bis 5 = immer), ergänzt durch Einzelitems zur Erfassung demografischer und berufsbezogener Merkmale. Die Erhebung der allgemeinen Selbstwirksamkeitserwartung erfolgte durch eine Skala von Schwarzer und Jerusalem

(1999) (10 Items, ω=.886). Es wurden latente Strukturgleichungsanalysen mithilfe des R-Packages lavaan durchgeführt.

Ergebnisse

Die konfirmatorische Analyse der Faktorenstruktur des TPD wies auf die Notwendigkeit einer Modifikation der bei Evers et al. (2016) gefundenen Dimensionen hin (X2(179)=431.8, p<.001, RMSEA=.064, SRMR=.064, CFI=.892). Auf Grundlage einer Analyse hinsichtlich fehlerhafter Modellspezifikationen (Saris et al., 2009) wurden drei doppelt ladende Items entfernt. Weiterhin ergab sich die Notwendigkeit, die auf das Lesen bezogene Dimension hinsichtlich der Rezeption von Printmedien und dem Informieren im Internet zu unterteilen. Das finale Modell mit den Dimensionen ‚Reflexion und Feedback' (M=3.45, SD=0.64), ‚Online Informieren' (M=3.40, SD=0.77), ‚Kooperation zur Unterrichtsentwicklung' (M=3.16, SD=0.77), ‚Erproben' (M=3.01, SD=0.62), Kooperation zur Schulentwicklung' (M=3.00, SD=0.72) und ‚Printmedien lesen' (M=2.91, SD=0.85) weist eine gute Passung auf (X2(118)=212.7, p<.001, RMSEA=.048, SRMR=.046, CFI=.951).

In der anschließenden Regressionsanalyse zeigten sich positive Gender-Effekte (Geschlecht: weiblich) auf die Dimensionen Erproben (β=.17**, SE=.06), Reflektieren/Feedback (β=.26***, SE=.07) und die Kooperationsformen (Schulentwicklung: β=.15*, SE=.06; Unterrichtsentwicklung: β=.23***, SE=.06). Bezogen auf das Alter ergab sich für das Lesen von Printmedien ein quadratischer Verlauf, der auf eine häufigere Rezeption bis etwa 50 Jahre mit einem anschließenden Rückgang hinweist (β=-.14*, SE=.06). Quer- und Seiteneinsteigende berichten, sich häufiger online zu informieren (β=.13*, SE=.06). Eine höhere allgemeine Selbstwirksamkeitserwartung weist positive Effekte auf das Erproben (β=.21**, SE=.07) und die schulentwicklungsbezogene Kooperation (β=.19***, SE=.07) auf.

Diskussion

Es zeigt sich, dass eine deutsche Adaption des TPD mit einigen Anpassungen im deutschsprachigen Raum Anwendung finden kann. Insbesondere die Nutzung von Online-Informationsquellen scheint sich von der bisherig verwendeten Dimension der allgemeinen unterrichtsbezogenen Informationsbeschaffung abzulösen. Die Studie ermöglicht einen Einblick in die Nutzungshäufigkeit von informellen Lerngelegenheiten in Deutschland, wonach Lehrpersonen häufig ihren Unterricht reflektieren, Feedback einholen sowie Online-Ressourcen zu bevorzugen zu scheinen. Limitierend ist zu berücksichtigen, dass es sich nur um Selbstberichte der Lehrpersonen handelt. Weiterhin zeigen sich bedeutsame, wenngleich kleine Effekte verschiedener demografischer, dispositioneller und berufsbezogener Merkmale, die im Rahmen des Vortrages diskutiert werden.

Lehrkräfte auf Instagram. Ein Ort für Austausch und soziale Unterstützung?

Eric Richter[1], Jeffrey P. Carpenter[2], André Meyer[1], Dirk Richter[1]
[1]Universität Potsdam, Germany, [2]Elon University, USA

Theoretischer Hintergrund

Die Zusammenarbeit mit Kolleg*innen ist für Lehrkräfte eine wichtige Quelle sozialer Unterstützung, indem sie zur Verminderung beruflicher Isolation und zum gemeinsamen Lernen beiträgt (de Jong et al., 2019; Vangrieken et al., 2015). Während Lehrkräfte bislang vor allem innerhalb der eigenen Schule zusammenarbeiteten, nutzen sie zunehmend auch soziale Medien (z.B. Twitter) als Ort für

den Austausch (Greenhow et al., 2020). Während einige Lehrkräfte hierbei auf niedrigschwellige Kooperationsformen zurückgreifen (z.B. durch den Austausch von Materialien), engagieren sich andere Lehrkräfte in intensiveren Kooperationsformen, etwa indem sie gemeinsam Unterrichtsmaterialien entwickeln (Greenhow et al., 2021). Bislang ist ungeklärt, ob sich verschiedene Kooperationsformen von Lehrkräften in sozialen Medien mithilfe standardisierter Messinstrumente reliabel erfassen lassen. Fener ist unklar, inwiefern verschiedene Kooperationsformen mit dem Erleben sozialer Unterstützung zusammenhängen. Hieraus ergeben sich zwei Forschungsfragen:

1. Können verschiedene Formen der Kooperation von Lehrkräften mithilfe eines standardisierten Messinstrumentes gemessen werden? Inwieweit berichten Lehrkräfte über unterschiedliche Formen der Zusammenarbeit auf Instagram?

2. Welche Zusammenhänge bestehen zwischen unterschiedlichen Formen der Lehrkräftekooperation und dem Erleben sozialer Unterstützung auf Instagram?

Methode

Bei der vorliegenden Studie handelt es sich um eine querschnittliche Fragebogenstudie. Sie basiert auf Daten von insgesamt 249 Lehrkräften, die Instagram im Rahmen ihrer beruflichen Tätigkeit nutzen.

Der Fragebogen umfasste Angaben zur beruflichen Nutzung von Instagram sowie zum Erleben sozialer Unterstützung auf Instagram. Hinsichtlich der beruflichen Nutzung von Instagram entwickelten wir in Anlehnung an Conze et al. (2020) insgesamt zehn Items zur Erfassung der drei Kooperationsformen Suche nach Informationen, Bereitstellen von Informationen und Ko-Konstruktion von Informationen. Die Items wurden auf einer sechsstufigen Likert-Skala eingeschätzt, die von 1 (nie) bis 6 (mehrmals täglich) reichte.

Zur Erfassung der wahrgenommenen sozialen Unterstützung auf Instagram nutzten wir zwölf Items aus der COACTIV-R-Studie und passten diese dem Untersuchungskontext an. Die Items bildeten die drei Subskalen emotionale Unterstützung ($\omega = 0.86$), informationelle Unterstützung ($\omega = 0.93$) und instrumentelle Unterstützung ($\omega = 0.74$) und wurden auf einer vierstufigen Likert-Skala eingeschätzt.

Zur Beantwortung der ersten Forschungsfrage wurden eine konfirmatorische Faktorenanalyse (CFA) sowie deskriptive Analysen durchgeführt. Zur Beantwortung der zweiten Forschungsfrage wurde ein Strukturgleichungsmodell (SEM) geschätzt, in dem die verschiedenen Formen der Kooperation als unabhängige Variablen und die Formen wahrgenommener sozialer Unterstützung als abhängige Variablen fungierten.

Ergebnisse

Die Ergebnisse der CFA deuten darauf hin, dass die Daten das postulierte Drei-Faktoren-Modell gut abbilden ($\chi2 = 67.39$, $df = 32$, $p = .01$, $CFI = .94$, $RMSEA = .07$, $SRMR = .05$). Die Mittelwerte der drei Skalen verdeutlichen, dass Lehrkräfte Instagram häufiger für niedrigschwellige Formen der Kooperationen, wie der Suche nach Informationen ($M = 3.78$, $SD = 1.03$) oder dem Bereitstellen von Informationen ($M = 2.12$, $SD = 1.01$), nutzen und seltener für intensive Formen wie der Ko-Kreation von Informationen ($M = 1.34$, $SD = 0.61$).

Die Befunde des SEM zeigen, dass die Zeit, die Lehrkräfte mit der Suche nach Informationen auf Instagram verbringen, negativ mit dem Erleben informationeller Unterstützung (β = -.19) und positiv mit dem Erleben instrumenteller Unterstützung zusammenhängt (β = .37). Das Bereitstellen von Information hängt dagegen positiv mit dem Erleben emotionaler (β = .56) und informationeller Unterstützung (β = .57) zusammen. Ko-Kreation von Information hängt zuletzt positiv mit instrumenteller Unterstützung zusammen (β = .25).

Relevanz

Die Studie verdeutlicht, dass Instagram für Lehrkräfte ein Ort des Austausches sein kann, der das Erleben sozialer Unterstützung ermöglicht. Gleichwohl machen die Befunde deutlich, dass Lehrkräfte auf Instagram insgesamt auf einem niedrigen Niveau miteinander kooperieren und dass mehr Zeit, die für Kooperation aufgewendet wird, nicht per se zu einer höheren wahrgenommenen sozialen Unterstützung führt. Dies könnte darauf hindeuten, dass Lehrkräfte spezifische digitale Kompetenzen benötigen, um die Potentiale sozialer Medien wie Instagram für die eigene berufliche Entwicklung vollumfänglich auszuschöpfen.

Investigating how Teachers Share and Use Materials in German Twitter Communities

Fitore Morina[1], Tim Fütterer[1], Joshua Rosenberg[2], Jeffrey Carpenter[3], Christian Fischer[1]
[1]Universität Tübingen, [2]University of Tennessee, Knoxville, United States, [3]Elon University, United States

Teachers use various materials and resources in their teaching (Brown, 2009; Charalambous & Hill, 2012). As textbooks may become outdated, many teachers access their professional network to receive more up-to-date and individualized teaching materials (Macià & García, 2016). In recent years, many teachers turned to social media to acquire teaching materials and to participate in informal learning (Greenhow et al., 2020). In particular, the micro-blogging platform Twitter is popular among teachers to get materials, ideas, and feedback for their classroom practice (Carpenter et al., 2015; Fischer, 2019; Rosenberg et al., 2020). Early research indicated that a primary reason for teachers to use Twitter is to acquire materials (Carpenter & Krutka, 2014, 2015). However, factors that contribute to how shared materials are used and perceived have not yet been examined at scale.

Research Questions

This project examines the sharing and usage of teaching materials in popular German Twitter communities using the following research questions (RQs). This abstract presents results for RQ1, whereas the presentation (if accepted) will share results for all RQs.

(RQ1) What types of materials are shared in German Twitter communities?

(RQ2) How are materials perceived and used in German Twitter communities?

(RQ3) Which user characteristics predict how materials are perceived and used?

Methods

Data were downloaded through the Twitter API with customized R and Python scripts. This project focused on four types of communities that are formed by combining all related individual hashtag communities within; (a) STEM, (b) Digital Learning, (c) the Twitterlehrerzimmer (TWLZ; German for “Teacher’s Twitter lounge,” Germany’s largest teacher community; Fütterer et al., 2021), and the chat-based community EDchatDE.

Tweet level data includes the tweet text, whether the tweet contains materials (any links and/or media; dichotomous), and continuous variables indicating the number of likes, retweets, and replies. A dichotomous variable indicates the tweet sentiment (positive/neutral vs. negative). User level data includes a dichotomous user classification (teacher/non-teacher) and continuous variables describing the number of followers and activity.

For RQ1, we calculated the percentage of tweets that included materials. For RQ2, we intend to use sentiment analysis to identify perception of materials and keyword analysis to examine if materials are implemented in classrooms. For RQ3, we will conduct ordinal regressions analyses with user characteristics as the dependent variable and tweet engagement as the independent variable.

Results

We found that 58.4% of tweets of the Digital Learning communities, of which 11.2% were posted by teachers, contain materials. In the STEM communities, 46.6% of tweets include shared materials, of which 17.7% were posted by teachers. In the TWLZ communities 29.9% of tweets include materials, of which 50.3% were posted by teachers. Lastly, in the EDchatDE communities only 29.6% of tweets, of which 32.1% were posted by teachers, contain materials.

Photos represented the highest percentage of shared media across community types (25.7% [EdchatDE] - 46.0% [TWLZ]). Videos and GIFs were posted much less, below 3% for all community types except the TWLZ communities, which included 9.4% tweets with GIFs and 11.8% tweets with videos.

Discussion

These results indicate that material sharing does not occur equally across community types. The lower percentage of materials associated with TWLZ and EDchatDE might indicate that these communities rather function as places of connection and emotional support (Fütterer et al. 2021; Rehm & Notten, 2016). Especially, the higher proportion of teachers posting materials might suggest more peer-to-peer interaction, which is beneficial for informal learning (Richter et al., 2014). In contrast, the STEM and Digital Learning communities are more specific in their topics, which might encourage teachers to use them for sharing and acquiring materials.

H-20: Feedback als Faktor für Lernerfolg und Unterrichtsentwicklung – welche Befunde liefern Interventionsstudien?

Zeit: Donnerstag, 02.03.2023: 15:30 - 17:15 · *Ort:* R11 T00 D05

Feedback als Faktor für Lernerfolg und Unterrichtsentwicklung – welche Befunde liefern Interventionsstudien?

Chair(s): **Mathias Ropohl** (Universität Duisburg-Essen, Germany), **Kerstin Göbel** (Universität Duisburg-Essen)
*Diskutant*in(nen):* **Marc Kleinknecht** (Leuphana Universität Lüneburg)

Feedback von Lehrkräften an ihre Schüler:innen gilt als ein Element von Unterricht, das einen hohen Effekt auf den Lernerfolg der Schüler:innen und damit auf die Wirkung von Unterricht hat. Seitdem Hattie (2020) in einer Metaanalyse den höchsten Effekt für das Unterrichtsmerkmal Feedback ausgewiesen hat, wird diesem sowohl in der Unterrichtspraxis als auch in der Bildungsforschung eine hohe Bedeutung beigemessen. Dennoch wird Feedback selten als Qualitätsmerkmal im Unterricht inkludiert. Wenn die Schüler:innen Feedback erhalten, erfüllt dieses auch nicht zwingend Qualitätsstandards für wirksames Feedback. Daher muss die Kompetenz des Feedback-Gebens und -Nehmens im Kontext der Lehrpersonenausbildung adressiert werden (Kleinknecht & Gröschner, 2016).

Insofern erscheint es wichtig, das Thema Feedback nicht nur für die Schule, sondern auch für die Lehrpersonenausbildung weiter in den Blick zu nehmen, beide Perspektiven zu verschränken und differentielle Effekte zu betrachten. Ein starker Fokus von Forschung muss die Analyse von Wirkzusammenhängen zwischen Feedbackmerkmalen und Personenmerkmalen sein.

Vor diesem Hintergrund umfasst das Vortragssymposium vier (quasi-)experimentelle Studien, die aktuelle Desiderata in Bezug auf die Forschung zu Feedback aufgreifen.

Der erste Beitrag fokussiert auf die Merkmale von Schüler:innen und deren Rolle beim feedbackgestützten Lernen. Als Merkmal wird die Mehrsprachigkeit in den Blick genommen. Damit bietet diese Untersuchung das Potenzial, den Einfluss von Mehrsprachigkeit auf die Nutzung und Verarbeitung von Feedback im Fach Deutsch zu analysieren.

Mit Blick auf die Qualität und Wirksamkeit von Feedback greift der zweite Beitrag kriteriengeleitetes Feedback im Englischunterricht als Thema auf. Damit wird dem Argument Rechnung getragen, dass Feedback häufig nicht sachbezogen ist und zudem für die feedbackgebenden Lehrkräfte einen hohen Zeitaufwand bedeutet. Kriterienraster bieten hier das Potenzial, das Feedback selbst transparent zu gestalten, aber auch die Lehrkräfte im Feedbackgeben zu unterstützen.

Der dritte Beitrag stellt eine Interventionsstudie im Kontext Lehrkräftebildung vor, im Rahmen derer Lehramtsstudierende kooperativ über eigene und fremde Unterrichtsvideos reflektieren und sich gegenseitig Feedback zum Unterricht geben, um eine hohe Reflexionsqualität zu erzielen. Der Beitrag verdeutlicht die hohe Relevanz von Feedback und kooperativen Videoanalysen für die Entwicklung von Feedbackkompetenzen von Lehramtsstudierenden.

Im vierten Beitrag werden Befunde aus einer Untersuchung vorgestellt, die die Wirkung von Feedback in der Lehrkräftebildung analysiert. Mit Hilfe des Feedbacks soll die Qualität von Lehr-Lern-Materialien für den Chemieunterricht erhöht werden und damit gleichzeitig potentiell die Unterrichtsqualität.

In allen vier Beiträgen zeigt sich, dass Feedback und insbesondere verschiedene Formen von Feedback in ganz unterschiedlichen Kontexten zu differenziellen Effekten führen können. Die Vortragenden werden spezifisch für den individuellen Schwerpunkt Perspektiven für weitere Forschungsarbeiten aufzeigen.

Beiträge des Symposiums

Qualität statt Quantität? Zum Zusammenhang von Feedback und Textqualität unter Berücksichtigung von migrationsbedingter Mehrsprachigkeit

Nora Müller, Till Utesch, Vera Busse
Westfälische Wilhelmsuniversität Münster

Theoretischer Hintergrund

Das Verfassen von Texten ist für viele Lernende herausfordernd, wobei herkunfts- und migrationsbezogene Disparitäten gerade im Schreiben besonders deutlich zutage treten (vgl. Autor 2022). Feedback kann insbesondere mehrsprachige Schreibende wirksam unterstützen (Biber et al. 2011). Dabei sollte lernförderliches Feedback Informationen zu Lernzielen beinhalten (Feed Up), Informationen über den Ist-Zustand im Vergleich zum Soll-Zustand (Feed Back) enthalten und Wege zur weiteren Verbesserung (Feed Forward) aufzeigen (Hattie & Timperley, 2007). Speziell in Bezug auf die Zusammenhänge zwischen diesen Feedbackkomponenten und Schreibleistungen mangelt es jedoch an Studien. Aktuelle internationale Befunde (Brooks et al., 2021) weisen zwar auf signifikante Zusammenhänge hin, berücksichtigen aber keine herkunftsbezogenen Merkmale der Lernenden.

Fragestellung

Inwieweit besteht ein Zusammenhang zwischen Feedbackpraktiken und Textqualität im Deutschen unter Berücksichtigung des Migrationshintergrunds und weiterer herkunftsbezogener Merkmale?

Methode

Die Daten stammen aus einer Interventionsstudie, in deren Rahmen Lehrkräfte zu diversitätssensiblem Feedback fortgebildet wurden. Für die Analyse werden Daten von einer Stichprobe aus N = 208 Lernenden in Klasse 9 an nichtgymnasialen Schulformen (MAlter = 14,03, SD = 0,75; n = 91 weiblich, n = 112 männlich; Migrationshintergrund n = 106) genutzt, die vor dem ersten Messzeitpunkt vor der Fortbildung erhoben wurden.

Neben der Erhebung von personen- und herkunftsbezogenen Merkmalen wurden Daten zur Quantität und wahrgenommenen Qualität von unterschiedlichen Feedbackmethoden zum Schreiben erfasst ($\alpha > 0{,}70$). Zur Erhebung der Textqualität wurden zwei Lernendentexte in Bezug auf Struktur und Kohärenz analysiert (Autoren, in Überarbeitung).

Ergebnisse und ihre Bedeutung

Es zeigen sich signifikante Korrelationen zwischen Textqualität und Feedbackqualität sowie dem Feed Forward zur Textentwicklung, nicht jedoch zwischen Textqualität und Feedbackquantität. Mehrebenen-Regressionsanalysen zeigen, dass die allgemeine Feedbackqualität und Feed Forward zur Textentwicklung unter gegenseitiger Kontrolle 23,5 % der Varianz in der Textqualität der Gesamtkohorte erklären (Modell 1). Es zeigt sich je ein signifikanter kleiner Effekt der

Feedbackqualität (ß = 0,17; p = 0,028) und des Feed Forwards zur Textentwicklung (ß = 0,16; p = 0,031) auf die Schreibleistungen der Gesamtstichprobe.

Nach Ergänzung des Migrationshintergrunds sowie Interaktionstermen (Modell 2) wurde 25,1 % der Varianz erklärt, wobei der Migrationshintergrund einen signifikanten negativen Effekt auf Unterschiede in der Textqualität hatte (ß = -0,20; p = 0,005); darüber hinaus zeigte sich ein signifikanter Effekt der Feedbackqualität (ß = 0,22; p = 0,004), die für beide Lernendengruppen einen signifikanten Beitrag zur Aufklärung von Unterschieden in der Textqualität leistete.

Die Ergänzung weiterer individueller Merkmale der Lernenden als Kontrollvariablen (Modell 3) erklärte 25,5 % der Varianz in der Textqualität. Die Feedbackqualität hatte auch hier einen signifikanten Effekt für alle Lernenden (ß = 0,25; p = 0,004). Weiterhin ergab sich eine signifikante negative Interaktion des Feed Forwards zur Textentwicklung mit dem Migrationshintergrund (ß = -0,17; p = 0,047). Simple-Slope-Analysen zeigten signifikante Unterschiede der beiden Steigungen von Lernenden mit und ohne Migrationshintergrund. Während also das Feed Forward zur Textentwicklung für Lernende ohne Migrationshintergrund (b = 0,477, p = 0,047) einen signifikanten Beitrag zur Aufklärung von Unterschieden in der Textqualität leistete, war dies bei Lernenden mit Migrationshintergrund nicht der Fall (b = -0,042).

Die Ergebnisse zeigen, dass wahrgenommene allgemeine Feedbackqualität und Feed Forward (also spezielle Hinweise zur Textentwicklung) einen signifikanten Anteil an Varianz der Textqualität vorhersagen; die Ergänzung der Feedbackhäufigkeit verbesserte das Modell nicht. Die Ergebnisse unterstreichen die Relevanz von Feedback für die Textqualität, wobei diskutiert werden kann, inwieweit die allgemeine Feedbackqualität wichtiger ist als die Feedbackquantität. Simple-Slope-Analysen legen nahe, dass Lernende mit Migrationshintergrund von Feed Forward zur Textentwicklung nicht profitieren. Es wird vermutet, dass das Feed Forward passgenauer an ihre Bedürfnisse angepasst werden muss. Der Vortrag diskutiert die Ergebnisse auch im Zusammenhang mit Daten zur Textqualität im Englischen sowie zum zweiten Messzeitpunkt, die derzeit ausgewertet werden.

Wie wirken Rubrics als Feedbackinstrument beim englischen Schreiben?

Stefan Daniel Keller[1], Ruth Trüb[2], Johanna Fleckenstein[3], Jennifer Meyer[4], Thorben Jansen[4]
[1]PH Zürich, [2]PH FHNW, [3]Universität Hildesheim, [4]Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Theoretischer Hintergrund

Feedback ist ein wirksamer Einflussfaktor auf das Lernen von Schülerinnen und Schülern (Hattie, 2020). Für das Feedback beim Schreiben fanden Graham et al. (2015) in einer Metaanalyse je nach Quelle der Rückmeldung (Lehrpersonen, Peers, Computer) Effektstärken zwischen d = 0.38 und d = 0.87. Trotz dieser positiven Befunde wird insbesondere prozessorientiertes Feedback von Lehrpersonen selten eingesetzt, was u.a. auf den hohen Aufwand zurückzuführen ist (Graham et al., 2015).

Feedback wirkt nur dann positiv auf die Leistung und Motivation von Lernenden, wenn es die Lücke zwischen Ist- und Sollzustand verdeutlicht und dadurch Unsicherheit und kognitive Belastung reduziert (Shute, 2008). Das Feedback-Modell von Hattie und Timperley (2007) fasst diese empirisch

identifizierten Wirksamkeitskriterien anhand von drei Fragen zusammen: „Wohin gehe ich?“ (Feed Up-Transparenz der Lernziele), „Wie komme ich voran?“ (Feed Back-Einschätzung des aktuellen Lernstands) und „Wohin geht es als nächstes?“ (Feed Forward-individuelle Hinweise zur Aufgabenerfüllung).

Beim Schreiben in den Fremdsprachen wird Feedback oft über Beurteilungsraster (Rubrics) umgesetzt. Dabei werden skalierte Leistungsniveaus für Qualitätsdimensionen hinsichtlich einer Aufgabe verwendet, um Lernziele explizit zu machen, eine Grundlage für die Einschätzung des aktuellen Lernstands zu liefern und aufzeigen, wie ein höheres Leistungsniveau erreicht werden kann (Allen & Tanner, 2006). Allerdings gibt es für den deutschsprachigen Raum kaum empirische Forschung, wie Rubrics als Feedbackinstrumente beim englischen Schreiben wirken.

Forschungsfragen

1) Kann rubric-basiertes Feedback die Schreibfähigkeit von Lernenden bei einer englischen E-Mail verbessern?

2) In welchen Bereichen verbessert sich die genre-spezifische Textqualität durch rubric-basiertes Feedback?

Methode

In einer computerbasierten Intervention schrieben N = 388 Lernende der 8. Klasse formale englische E-Mails zu alltagsrelevanten Situationen (z.B. Informationen zu einem Sprachkurs in England erfragen). Insgesamt wurden E-Mails zu drei Aufgaben mit vergleichbarer Schwierigkeit geschrieben. Im Anschluss an jede E-Mail wurde eine Rückmeldung anhand eines Rubric generiert. Dieses hatte fünf aufeinander aufbauende Leistungsstufen, die sich auf genrespezifische Aspekte einer formalen E-Mail bezogen. In Anlehnung an das Modell von Hattie und Timperley (2007) wurden drei Feedback-Bedingungen untersucht, bei denen das Rubric in Kombination mit unterschiedlichem Informationsgrad verwendet wurden: Interventionsgruppe (IG) 1 erhielt das Rubric ohne zusätzliche Informationen; IG 2 erhielt das Rubric zusammen mit einer Einschätzung der Textqualität, und IG 3 erhielt das Rubric zusammen mit einem kurzen Text, der Informationen darüber enthielt, wie sie ihre Leistung verbessern konnten. Eine Kontrollgruppe (KG) erhielt alle Rückmeldungen erst nach Abschluss der Studie. Sowohl die drei IGs als auch die KG erhielten einen „Schreibrahmen“, welcher zentrale Elemente von Emails enthielt, auf die sich die Lernenden beim Schreiben beziehen konnten. Die abhängige Variable war die Textqualität der Schülertexte. Diese wurde nach einem generischen Analyseschema von jeweils zwei trainierten Rater:innen ausgewertet (Kendall’s W .88-.97).

Ergebnisse und ihre Bedeutung

Es zeigt sich in allen Gruppen eine signifikante Verbesserung der genre-spezifischen Textqualität zwischen den Aufgaben 1 und 2 ($d = 0.94$, $p < .001$) und 1 und 3 ($d = 0.93$, $p < .001$). Jedoch konnten keine signifikanten Unterschiede zwischen den Aufgaben 2 und 3 gezeigt werden, und zwar weder zwischen IGs und KG, noch zwischen den verschiedenen IGs. Es kann vermutet werden, dass der „Schreibrahmen“ stark gewirkt und differenzielle Effekte des Feedbacks zugedeckt haben könnte.

Bei einer linguistischen Detailanalyse der Textqualität in den IGs zeigte sich, dass insbesondere jene Textaspekte von der Rückmeldung profitieren, welche durch genre-typische Ausdrucksformeln gelöst werden können (z.B. salutation, concluding sentence). Hier unterschied sich die Textqualität in der ersten und zweiten Schreibaufgaben hochsignifikant voneinander. Textaspekte, welche einen freieren Sprachgebrauch voraussetzen, wurden weniger stark beeinflusst.

In nachgelagerten Entwicklungsarbeiten ist es der Forschungsgruppe gelungen, eine KI-basierte Feedback-Engine zu entwickeln, mit welcher sich text- und genrespezifisches Feedback automatisiert erstellen lässt. Damit lassen sich Feedbackeffekte in künftigen Studien differenzierter untersuchen.

Kooperative Unterrichtsvideoreflexion in der Lehrkräftebildung – wie entwickeln sich feedbackbezogene Einstellungen und Reflexionsqualität angehender Lehrkräfte?

Kerstin Göbel, Julia Bönte
Universität Duisburg-Essen

Theoretischer Hintergrund

Im Diskurs um die Professionalisierung von Lehrkräften in Deutschland wird der Ausbildung der Reflexionskompetenz große Bedeutung beigemessen (Te Poel & Heinrich, 2020). Kooperative Reflexionsformate scheinen dabei besonders vielversprechend (Brown et al., 2021), da sie eine Perspektiverweiterung und den Austausch neuer Perspektiven ermöglichen (Körkkö, 2019; Wyss, 2013). Lernmöglichkeiten werden sowohl durch das Geben als auch durch das Empfangen von Feedback geschaffen (Adachi, Tai & Dawson 2018; Narciss, 2006).

Da viele relevante Aspekte einer Unterrichtssituation mittels Videografie dokumentiert werden können, eignen sich Videos in besonderem Maße für eine systematische und vertiefende Reflexion des Unterrichts (Thissen et al., 2019) und werden zunehmend als Lern- und Reflexionsmedium eingesetzt (z. B. Weber et al., 2018). Videogestützte und kooperative Reflexionsformate erweisen sich auch für die Kompetenzentwicklung von (angehenden) Lehrkräften als förderlich (Krammer et al., 2015; Seidel & Thiel, 2017). Während eigene Videos insbesondere eine hohe emotionale Involviertheit und Motivation begünstigen (Richey et al., 2017), ermöglicht die Reflexion anhand fremder Videos eine kritische, distanziertere Auseinandersetzung (z. B. Kleinknecht & Schneider, 2013). Die Qualität von Reflexion und Feedback im Rahmen von kooperativen videobasierten Reflexionssettings (mit Peers) wurde bislang kaum untersucht.

Forschungsfragen

1. Führt die videogestützte kooperative Unterrichtsreflexion in einem strukturierten Reflexionssetting anhand eigener und fremder Videos zur Verbesserung der Einstellungen gegenüber Feedback und videogestützter Reflexion?

2. Welche Reflexions- und Feedbackqualität kann im Rahmen der strukturierten kooperativen Unterrichtsvideoreflexion erreicht werden?

Methode

In Rahmen der Studie reflektierten Lehramtsstudierende im Praxissemester anhand von fremden (IG1, n = 36) und eigenen (IG2, n = 21) Unterrichtsvideos in einem strukturierten kooperativen Reflexionssetting (Autoren, 2018). Durch die gezielte (Vor-)Strukturierung und Unterstützung des

Reflexionsprozesses in Anlehnung an das Interactive-Two-Feedback-Loops-Modell [IFTL] nach Narciss (2006) sollten die Bedingungen für die Effizienz von Feedback begünstigt werden, u.a. durch präzise Bestimmung der Aufgabenanforderungen, die Qualität der Informationsverarbeitung der lernenden Person (interner Feedback-Loop) und der Peers (externer Feedback-Loop). Als Kontrollgruppe (KG, n = 26) diente ein Seminar ohne derartiges Reflexionssetting.

Im quantitativen Teil der Studie (quasi-experimentelles Prä-Post-Kontrollgruppendesign) wurden u. a. die Einstellungen gegenüber Feedback (Behnke, 2016) und die Akzeptanz von videogestützter Unterrichtsreflexion (Kleinknecht, Ottinger & Schneider, 2014) erfasst. Die Datenanalyse zum Vergleich von Experimentalgruppen und Kontrollgruppe erfolgte mittels Varianzanalysen.

Im Rahmen einer qualitativen Inhaltsanalyse (n = 5 Lehramtsstudierende) wurde explorativ die kooperative Unterrichtsvideoreflexion im zuvor videografierten strukturierten Reflexionssetting basierend auf dem ERTO-Reflexionsmodell im Hinblick auf die Reflexions- und Feedbackqualität vertiefend analysiert (Krieg & Kreis, 2014). Nach Transkription von ca. 90 Minuten Videomaterial aus drei Reflexionsphasen erfolgte die kategoriengeleitete Datenauswertung mittels qualitativer Inhaltsanalyse nach Mayring (2010).

Ergebnisse und ihre Bedeutung

Die Analyseergebnisse verdeutlichen gruppenübergreifend eine Zunahme der empfundenen Bedeutung von Feedback im Allgemeinen (Zeiteffekt: $F(1.67) = 5.325$, $p = 0.024$; $\eta 2 = 0.074$), wobei sich gruppenspezifisch lediglich in IG2 (Reflexion anhand eines eigenen Videos) signifikante Einstellungsverbesserungen zeigen. Es zeigen sich weiterhin signifikante Unterschiede zwischen den Interventionsgruppen bezüglich der Einstellungsveränderung gegenüber der videogestützten Unterrichtsreflexion zugunsten von IG2.

In Hinblick auf die Reflexions- und Feedbackqualität zeigt sich, dass die Studierenden bereits zu einer vertieften Auseinandersetzung mit dem Reflexionsgegenstand fähig sind (introspektive Reflexion und Feedback; Narciss, 2006, 2017). Die Qualität der Reflexion im internen Feedback-Loop (u. a. Überwachung des Lernprozesses durch die lernende Person), aber auch im externen Feedback-Loop (u. a. Qualität der Generierung) scheinen begünstigt worden zu sein. Allerdings fehlt im Feedbackprozess die Einbeziehung theoretisch und empirisch fundierter Argumente.

Die Befunde deuten einerseits auf die Relevanz des im Rahmen der Reflexion verwendeten Videotyps (eigenes Video) für die Einstellungsveränderungen angehender Lehrkräfte durch strukturiertes Peer-Feedback hin. Sie liefern zudem explorative Erkenntnisse zur Möglichkeit der Erreichung einer hohen Reflexionsqualität durch kooperatives Peer-Feedback und bieten darüber hinaus Anknüpfungspunkte, an denen Optimierungsstrategien für videogestützte Peer-Reflexion ansetzen können.

Welches Feedback wirkt am besten? – Wirksamkeit von externem und internem Feedback bei Lehramtsanwärter:innen

Büsra Tonyali, Mathias Ropohl
Universität Duisburg-Essen

Theoretischer Hintergrund

Im naturwissenschaftlichen Unterricht dienen multiple externe Repräsentationen als grundlegendes Medium der Wissensvermittlung (Ainsworth, 2006). Untersuchungen zeigen jedoch, dass angehende Chemielehrkräfte über ein geringes Wissen zu Repräsentationen verfügen (Taskin, Bernholt & Parchmann, 2015). Insbesondere stellen Wechsel zwischen Repräsentationsebenen eine Schwierigkeit dar (van Driel & Verloop, 1999). Diese Wechsel sind gerade in der Unterrichtspraxis relevant, wenn Lehr- und Lernmaterialien fachlich und fachdidaktisch akkurat gestaltet und eingesetzt werden sollen. Befunde legen nahe, dass Feedback das Potenzial hat, solche praxisbezogenen Kompetenzen von Lehrkräften während ihrer Ausbildung zu fördern (Lipowsky, 2009). Es wird angenommen, dass die Kombination aus internem (selbstgeneriertem) sowie externem (fremdgeneriertem) Feedback besonders wirksam ist (Butler & Winne, 1995). Vor diesem Hintergrund wurde als Ziel eine Unterstützungsmaßnahme entwickelt, die Lehramtsanwärter:innen (LAA) dabei helfen soll, ihre Kompetenzen in Bezug auf den Umgang mit Repräsentationen zu erweitern.

Fragestellung

Inwiefern lassen sich repräsentationsbezogenes Fachwissen (repFW) und fachdidaktisches Wissen (repFDW) mithilfe von internem und/oder externem Feedback bei LAA fördern?

Methode

Zur Beantwortung der Forschungsfrage wurde eine Interventionsstudie im Prä-Post-Kontrollgruppendesign als Online-Selbstlernmodul im Vorbereitungsdienst in NRW implementiert (N = 90 LAA). Um Interaktionseffekte zu untersuchen, wurde das Feedback als unabhängige Variable in den Untersuchungsgruppen systematisch variiert: 1. externes und internes (IG1), 2. internes (IG2), 3. kein Feedback (KG). Das interne und das externe Feedback wurden anhand eines Bewertungsbogens standardisiert generiert.

Die Intervention bestand aus drei Sitzungen à 90 Minuten. Jede Sitzung war identisch aufgebaut und thematisierte einen anderen chemischen Inhalt. Im ersten Schritt designten die LAA ein Lehr-Lern-Material anhand eines vorgegebenen Lernziels sowie bereitgestellten Text- und Bildelementen. Anschließend bewertete jede Testperson in IG1 und IG2 ihr Arbeitsblatt mittels Bewertungsbogen (internes Feedback). In der folgenden Interventionssitzung erhielten die Testpersonen in IG1 einen für ihr Arbeitsblatt ausgefüllten Bewertungsbogen von Seiten der Testleitung (externes Feedback). Zuletzt verglichen sie beide Feedbacks miteinander und designten in der nächsten Sitzung erneut ein Lehr-Lern-Material. Die IG2 erhielt kein externes Feedback und die KG gar kein Feedback. In diesen beiden Gruppen wurden die Arbeitsphasen mit anderen nicht feedbackbezogenen Aufgaben ergänzt, um die Gesamtarbeitszeit in den Gruppen zu kontrollieren. Die Wirkung der Intervention wurde mit standardisierten Tests, die den Wissenszuwachs des repFW (α = .69) und repFDW (α = .63) erheben, überprüft.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen im Prätest, dass das repFW (M = 50-59 %) und repFDW (M = 36-43 %) theoriekonform und wie erwartet in einem geringen Umfang vorliegen. Eine Zunahme durch die Intervention liegt im Posttest beim repFW in der IG1 (d = .52, p = .009) sowie beim repFDW in der IG2 (d = .62, p = .002) vor. Beide Wissenszuwächse indizieren eine mittelgroße Effektstärke. Außerdem

liegen zu beiden Interventionsgruppen Daten zum Cognitive Load für alle drei Interventionssitzungen vor. Auffällig ist der Cognitive Load bzw. die wahrgenommene Schwierigkeit beim Ausfüllen des Bewertungsbogens. Diese fällt in beiden Gruppen IG1 und IG2 von der ersten Sitzung bis zur dritten Sitzung signifikant ab. Der Cognitive Load ist jedoch in IG1 in allen Sitzungen signifikant höher als im Vergleich zu IG2 ($d = .73$-$.92$, $p = .021$-$.050$).

Zusammenfassend deuten die Daten daraufhin, dass internes und/oder externes Feedback differenzielle Effekte bei angehenden Chemielehrkräften bewirken: Internes Feedback bewirkt die Zunahme von repFDW, während beim zusätzlichen Einsatz von externem Feedback das repFW, aber auch der Cognitive Load steigen. Bezogen auf den Kontext Vorbereitungsdienst, in dem fachliches Wissen der LAA aus dem Lehramtsstudium bereits ausgereift sein müsste und schwerpunktmäßig fachdidaktisches Wissen vertieft werden muss, erweist sich in dieser Studie internes Feedback daher als die effektivere Feedbackform.

In den nächsten Analyseschritten werden die Überzeugungen, die Qualitäten der designten Lehr-Lern-Materialien sowie die Qualitäten des internen Feedbacks ausgewertet. Im Vortrag werden die detaillierten Befunde berichtet und die Wirkung des Feedbacks kritisch diskutiert.

Postersession

Zeit: Dienstag, 28.02.2023: 19:00 - 20:30 · *Ort:* Mensa

1. EnvironMINT: Gelingensbedingungen für eine erfolgreiche Zusammenarbeit von Kindern, Eltern, Schulen und FabLabs für nachhaltige MINT-bezogene Maker-Aktivitäten

Kathrin Smolarczyk[1], Sophie Uhing[2], Stephan Kröner[2]
[1]Hochschule Rhein-Waal, Germany; [2]Friedrich-Alexander-Universität Erlangen-Nürnberg

Theoretischer Hintergrund des Projekts

Laut dem MINT-Aktionsplan gelingt es häufig nicht, die kindliche Neugier und Experimentierfreudigkeit für Naturwissenschaft und Technik über die Jugend hinweg aufrechtzuerhalten (Bundesministerium für Bildung und Forschung, 2019). Der Unterricht in MINT-Fächern steht dabei vor der Herausforderung, eine Verbindung von abstrakten MINT-Themen zu fächerübergreifenden Fragen des Alltags und praktischen Anwendungen herzustellen und Jugendliche nachhaltig zu begeistern. Um das Interesse an und die Motivation für MINT nachhaltig zu fördern, sollte nicht erst am Ende der Sekundarstufe I angesetzt werden (Chittum, Jones, Akalin & Schram, 2017). Außerdem sollte das Potential von non-formalen Lernsettings gezielt genutzt und insbesondere die Zusammenarbeit zwischen MINT-Bildungssettings und Familien intensiviert werden (Bundesministerium für Bildung und Forschung, 2019). Eine Keimzelle für kreativen, selbstbestimmten Umgang mit MINT-Inhalten bietet die Maker-Community, zum Beispiel in offenen Werkstätten wie FabLabs. Durch kreative Maker-Aktivitäten wird träges Wissen aus schulischem MINT-Unterricht in fächerübergreifenden Praxisprojekten angewandt und vertieft. Mit digitalen und analogen Materialien und Werkzeugen einschließlich 3D-Druckern, E-Textilien, Einplatinen-Computern und Handwerkzeug können Teilnehmende eigene Ideen selbstbestimmt verwirklichen. Im Sinne eines ökosystemischen Ansatzes (Bronfenbrenner, 1981) stehen FabLabs dabei an der Schnittstelle der Subsysteme, die durch Eltern, Peers und Schule bestimmt sind. Dass die Maker-Kultur MINT-Kompetenzen und Interesse im MINT-Bereich fördern kann, insbesondere bei Mädchen und Jugendlichen mit niedrigem sozioökonomischen Status (SES), wurde international bereits vielfach dargelegt (Bevan, 2017). Durch FabLabs können Herausforderungen des Zugangs für entsprechende Medien und Technologien für Familien mit niedrigem SES adressiert werden (Dickens, Jordan & Lande, 2016; Rajan, 2021). FabLabs werden daher als eine der wichtigsten Säulen im Bereich der Bildung zum Umgang mit Technik betrachtet (Kinnula, Sánchez Milara, Norouzi, Sharma & Iivari, 2021). Insbesondere bei der Betrachtung von non-formalen Bereichen wurde die Beteiligung von Eltern als aktive Mitwirkende als wertvolle Möglichkeit gesehen, die Nachhaltigkeit von MINT-Themen in den Familien zu steigern (Dickens et al., 2016; Letourneau, McMillan Culp & Wells, 2021).

Ziel und Fragestellungen des Projekts

Das Projekt EnvironMINT untersucht die Anforderungen und Gelingensbedingungen für eine erfolgreiche Zusammenarbeit von Schulen und FabLabs mit Kindern, Eltern und Peers für selbstbestimmtes, nachhaltiges und kreatives Making als eine Form ganzheitlicher MINT-Bildung. Ziel des Projekts ist die Entwicklung eines integrativen Konzepts zur ganzheitlichen MINT-Bildung an der Schnittstelle von FabLabs, Schule, Peergroup und Elternhaus. Dabei verfolgt es folgende Fragestellungen: Welche Anforderungen und Gelingensbedingungen sind aus Sicht der jeweiligen

Akteure entscheidend für eine erfolgreiche Zusammenarbeit von Schulen und FabLabs mit Jugendlichen, Eltern und Peers für selbstbestimmtes und kreatives Making als eine Form ganzheitlicher MINT-Bildung? Wie kann das Thema Nachhaltigkeit gewinnbringend in das Konzept integriert werden?

Methodischer Aufbau des Projekts

Um das komplexe Zusammenspiel aus formaler, informeller und non-formaler Bildung im Bereich Making zu vereinen, findet in einem partizipativen und iterativen Forschungs- und Entwicklungsprozess eine Synthese aus einer Design-Case-Study (Wulf, Rohde, Pipek & Stevens, 2011) sowie dem Design-Thinking (Brown, 2008) einerseits und klassischen Methoden der empirischen Bildungsforschung wie Interviewstudien und Feldexperimenten mit Prä-Post-Kontrollgruppendesign andererseits statt. Die Erkenntnisse werden in ein integratives und skalierbares Konzept überführt, das das Potenzial von Making für die MINT-Bildung mehrperspektivisch ausschöpft, den aktiven Lernenden im Mittelpunkt sieht und das Thema Nachhaltigkeit gewinnbringend integriert: Zunächst geben ergebnisoffene Vorstudien mit Kindern, Eltern, Lehrkräften und Makern erste Aufschlüsse zu den Anforderungen der Akteure sowie Vorstellungen und Wünsche der Teilnehmenden bzgl. Maker-Aktivitäten. Hier wird der Fokus insbesondere auf Mädchen und Kindern mit niedrigem SES gerichtet. Anschließend finden in der Gestaltungsphase die Zusammenführung der Vorstudien sowie partizipative Designaktivitäten mit den Akteuren statt, die in die Konzepterarbeitung einfließen. In der Aneignungsphase wird das erarbeitete Konzept in unterschiedlichen Kontexten schließlich mehrfach praktisch erprobt und evaluiert. Hier wird insbesondere auch die Übertragbarkeit auf unterschiedliche Kontexte exploriert, wie Schulen mit und ohne Anbindung an FabLabs und Makerspaces.

2. Individuelles und situatives Interesse im Schüler:innenlabor - eine Latent-State-Trait Analyse

Xenia Schäfer, Sebastian Habig
Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Theoretischer Hintergrund:

Interessensförderung ist ein Ziel außerschulischer MINT-Lernorte (Tillmann & Wegner, 2021), zumal Interesse mit Berufswahlentscheidungen zusammenhängt (Su et al., 2019), die gerade im MINT-Bereich einer ungünstigen Entwicklung unterliegen (Köller et al., 2021).

Gemäß der Person-Gegenstands-Konzeption (Krapp, 1992) handelt es sich bei Interesse um eine Beziehung zwischen einer Person und einem Interessensgegenstand (z. B. Objekt/Tätigkeit). Die komplexe Dynamik hinter dem Interessenskonstrukt ergibt sich aus der Wechselwirkung relativ stabiler Persönlichkeitsmerkmale (individuelles Interesse, Trait) und einer situationsspezifischen und folglich variierenden Zustands-Komponente (situatives Interesse, State) (Krapp, 1992). Die beiden Varianten schließen sich nicht gegenseitig aus, sondern interagieren miteinander (Su et al., 2019). Auch das Erwartung-mal-Wert-Modell für motivationsgeleitetes Handeln im akademischen Kontext wurde von Eccles & Wigfield (2020) im Hinblick auf die Situationsabhängigkeit der darin enthaltenen Konstrukte aktualisiert (situatives Erwartung-mal-Wert-Modell).

Laut dem Modell bestimmt das Wirkungsgefüge aus Personen- und Umweltvariablen die Bereitschaft

eines Individuums eine Tätigkeit auszuführen und lässt Aussagen bezüglich Wahlentscheidungen, Engagement und Persistenz aus motivationaler Perspektive zu (Rosenzweig et al., 2019). Die Variablen, unter denen sich auch das individuelle und das situative Interesse befinden, werden stets situationsspezifisch gewichtet und miteinander in Bezug gesetzt.

Erfolgt die Erhebung von Interessensdaten bei Schüler:innen innerhalb von naturwissenschaftlichen Lernarrangements, so spielt die Aktivitätsform eine größere Rolle als das zugrunde liegende Themengebiet (Swarat et al., 2012). Daher erscheint bei der Betrachtung von Interesse in solchen Lernumgebungen eine tätigkeitsbasierte Differenzierung nach dem RIASEC+N-Modell (Blankenburg & Sheersoi, 2018) sinnvoll.

Erkenntnisse darüber zu erlangen, wie sich die Wechselwirkungen der beiden Interessens-Varianten in den jeweiligen Tätigkeitsphasen äußern und die Anteile aufzuschlüsseln, die auf Person bzw. Situation zurückzuführen sind, könnte Ansatzpunkte liefern, um MINT-Lernangebote interessensförderlicher zu gestalten und weiterführend MINT-Berufsorientierung positiv zu beeinflussen (Tillmann & Wegner, 2021).

Fragestellungen:

FF1: Inwiefern verändert sich situatives Interesse in Abhängigkeit der Aktivitätsphase im außerschulischen Lernort Schüler:innenlabor?

FF2: Welchen Einfluss haben individuelle Lernvoraussetzungen auf die Ausprägungen des situativen Interesses in Abhängigkeit der Aktivitätsphase im außerschulischen Lernort Schüler:innenlabor?

Methodik:

Für die Untersuchung der Forschungsfragen wurde ein Schüler:innenlaborprogramm (Sekundarstufe I, Realschule, Chemie) entwickelt, welches nach dem RIASEC+N-Modell (Blankenburg & Sheersoi, 2018) strukturiert wurde. Die Tätigkeiten umfassen dabei die Durchführung eines Experiments (Realistic), die Interpretation von Versuchsergebnissen (Investigative), das Anfertigen einer Versuchsskizze (Artistic), das Notieren von Messergebnissen (Conventional), das Präsentieren von Versuchsergebnissen (Enterprising) und Diskussion (Networking). Um die Tätigkeitsbindung der quantitativen Messungen umzusetzen, wird ein iterativer, mikroanalytischer Ansatz verfolgt (Rotgans & Schmidt, 2018). Daher erfolgt die Datenerhebung mittels eines Kurzskalen-Fragebogens zur Erfassung des situativen Interesses (States) zu sieben Messzeitpunkten jeweils nach Durchführung der RIASEC+N-Aktivitäten. Die Wahl der Kurzskalen-Variante soll die Störung des Workflows und damit einhergehende unerwünschte Einflüsse auf die Messung reduzieren.

Um den Zusammenhang zu den individuellen Lernvoraussetzungen (Traits) betrachten zu können, wird dem Laborangebot ein domänenspezifischer Fragebogen (Fechner, 2009) zu individuellem Interesse, Selbstkonzept, extrinsischer/intrinsischer Lernmotivation, Selbstwirksamkeitserwartungen und demografischen Daten vorgeschaltet.

Für die Auswertung der Daten bietet sich neben der deskriptiven Herangehensweise eine Latent-State-Trait Analyse mithilfe von Strukturgleichungsmodellen an, um die Anteile aufzuschlüsseln, die auf Person bzw. Situation während der jeweiligen Labortätigkeit zurückzuführen sind.

Erste Ergebnisse und Ausblick:

Die Pilotierung (*N* = 136) deutet auf einen Interessensabfall während der Enterprising-Phase hin (*M* = 3.22, *SD* = 1.50, 7-stufige Likert-Skala). Nach Überarbeitung des Studiendesigns wird dieses Ergebnis bestätigt (*M* = 3.69, *SD* = 1.40 bei *N* = 19) und durch ein Minimum bei der Investigative-Phase (*M* = 3.75, *SD* = .93) sowie ein Maximum bei der Social-Tätigkeit (*M* = 5.24, *SD* = .88) ergänzt. Dies wird durch eine Verfeinerung der Messmethodik sowie Vergrößerung der Stichprobe genauer ausdifferenziert und diskutiert. Darauf basierend lassen sich Empfehlungen für das Design ähnlicher MINT-Förderungsangebote ableiten. Zudem können bis zur Tagung erste Ergebnisse der Latent-State-Trait Analyse vorgelegt werden.

3. Situationales Interesse bei mathematischen Modellierungsaufgaben: Erhöhung durch Experimente möglich?

Sebastian Geisler[1], **Stefanie Rach**[2]
[1]Universität Hildesheim, Deutschland; [2]Otto-von-Guericke-Universität Magdeburg

Modellieren gehört zu den zentralen mathematischen Kompetenzen (KMK, 2022). Gerade im Zuge der Corona-Pandemie wird deutlich, wie relevant der Umgang mit Modellen (z. B. des exponentiellen Wachstums und Zerfalls) ist. Modellieren bedeutet, ausgehend von einer Realsituation ein vereinfachtes reales Modell zu entwickeln und dieses in ein mathematisches Modell zu übersetzen. Am mathematischen Modell entwickelte Resultate werden mit Blick auf die Realsituation interpretiert, bevor der gesamte Prozess validiert wird.

Der Fachkräftemangel im MINT-Bereich macht die Notwendigkeit deutlich, dass Schüler*innen Interesse an MINT-Aktivitäten, wie dem Modellieren, entwickeln. Interesse definieren wir als besondere Beziehung zwischen einer Person und einem (abstrakten) Objekt mit emotionsbezogenen (bei der Beschäftigung mit dem Interessenobjekt werden positive Emotionen erlebt) und wertbezogenen (dem Interessenobjekt wird ein subjektiv hoher Wert zugeschrieben) Valenzen (vgl. Krapp, 2007). Während situationales Interesse einen zeitlich begrenzten, durch erhöhte Aufmerksamkeit gekennzeichneten Zustand als Reaktion auf eine als interessant wahrgenommene Situation (z. B. eine konkrete Aufgabe) bezeichnet, gilt individuelles Interesse als stabilere Persönlichkeitsdisposition. Bisherige Studien zeigen, dass Schüler*innen Modellierungsaufgaben als weniger interessant wahrnehmen als innermathematische Aufgaben (z. B. Krug & Schukajlow, 2013).

Ein Ansatz, um das situationale Interesse zu steigern, ist die Verbindung von Experimentieren und Modellierungsaufgaben. Grundgedanke ist, dass Schüler*innen naturwissenschaftliche Experimente durchführen und die dabei gewonnenen experimentellen Daten als Ausgangspunkt für mathematische Modellierungen nutzen. Bisherige Studien deuten darauf hin, dass Aufgaben, die das Durchführen von Experimenten und Modellierungsprozesse kombinieren, das situationale Interesse von Schüler*innen steigern können (Beumann, 2016). Jedoch ist beispielsweise die Studie von Beumann (2016) ohne Kontrollgruppe durchgeführt und es fehlen Erkenntnisse dazu, welche Bedingungen in der Kombination von Experiment und Modellierungsprozess zu höherem situationalen Interesse führen. Daraus ergeben sich die Forschungsfragen:

1. Inwiefern unterscheidet sich das situationale Interesse während der Bearbeitung von Modellierungsaufgaben ohne Experimente und der Bearbeitung von Modellierungsaufgaben mit Experimenten?

2. Welche Gründe führen zu hohem oder niedrigem situationalen Interesse während der Bearbeitung von Modellierungsaufgaben mit Experimenten?

Zur Beantwortung der Fragen wurden zwei Studien durchgeführt. In Studie 1 bearbeiteten 46 Schüler*innen aus zwei zehnten Klassen jeweils eine Modellierungsaufgabe mit und eine ohne Durchführung eines Experiments und gaben danach ihr situationales Interesse mittels eines Fragebogens an. Situationales Interesse wurde mit vier Items auf einer fünfstufigen Likert-Skala von „stimmt überhaupt nicht“ (1) bis „stimmt vollkommen“ (5) gemessen, z. B. „Diese Aufgabe finde ich interessant“ (vgl. Willems, 2011, Cronbachs α = .86). Die Modellierungsaufgaben waren der Bierschaumzerfall und das Abkühlen von Tee – die Reihenfolge der Aufgaben und der Position des Experiments waren zufällig. In Studie 2 wurden 12 Zehntklässler*innen nach dem Bearbeiten einer der beiden Modellierungsaufgaben mit Experiment aus Studie 1 in problemzentrierten Interviews zu ihrem Interesse befragt. Die Antworten wurden mittels induktiver, qualitativer Inhaltsanalyse nach Mayring (2010) ausgewertet – Intercoder-Reliabilität κ = .65.

Die Ergebnisse von Studie 1. zeigen, dass Schüler*innen Aufgaben mit Experimenten ein höheres situationales Interesse (d=0,32, kleiner bis mittlerer Effekt) zuschreiben.

In Studie 2 wurde identifiziert, welche Gründe das Auftreten oder Ausbleiben von situationalem Interesse bedingen. Auffällig ist, dass die Schüler*innen das Durchführen des Experimentes mit Freude verbinden und dabei insbesondere den Kontext „Bierschaumzerfall“ nennen. Dagegen ist das Durchführen des Modellierungsprozesses mit weniger Freude verbunden, auch weil diesem eine größere Schwierigkeit zugesprochen wird. Dem Experimentieren selber wird kein eigenständiger Wert zugesprochen, sondern nur in Verbindung mit der anschließenden Modellierung.

Die Ergebnisse zeigen, dass Schüler*innen Modellierungsaufgaben mit Experimenten mehr Interesse zuschreiben als ohne Experimenten. Gründe für diese Unterschiede liegen nach den Äußerungen der Schüler*innen in der Leichtigkeit des Experimentierens und den verwendeten Kontexten, was die Bedeutung von sinnvollen Kontexten fürs Modellieren zeigt (vgl. van Vorst et al., 2018 im Bereich Chemie). Die quantitativen Ergebnisse sollten in einer größeren Stichprobe repliziert werden und auch die Auswirkungen des Experimentierens auf die Modellierungskompetenzen untersucht werden.

4. Ablenkend oder anregend? Eine qualitative Analyse kognitiver und motivationaler Prozesse bei der Bearbeitung interessensbasierter Sachunterrichtsaufgaben in der Grundschule

Ann-Kathrin Laufs, Sebastian Kempert
Universität Potsdam, Germany

Theoretischer und empirischer Hintergrund

Die Berücksichtigung von individuellen Interessen im Unterricht ermöglicht motivations- und lernförderliche Zugänge in unterschiedlichen Lernbereichen (Krapp, 2010). Die Kontextpersonalisierung stellt eine Möglichkeit dar, die außerschulischen Interessen von Schüler*innen bei der Gestaltung von Lernaufgaben aufzugreifen, und somit individuelle Zugänge zu neuen Themen zu ermöglichen (Lohrmann, 2014; Walkington, 2013). Der Forschungsstand zu den Effekten einer interessensbasierten Kontextpersonalisierung ist uneinheitlich: Während positive Effekte auf das situationale Interesse in der Lernsituation und das individuelle Interesse am Lerngegenstand empirisch belegt sind (Bernacki & Walkington, 2018; Laufs & Kempert, 2021a;

Reber, Canning & Harackiewicz, 2018), zeigen sich hinsichtlich der Effekte auf den Lernerfolg uneinheitliche Befunde (Übersicht in: Walkington & Hayata, 2017). Zur Erklärung dieser uneinheitlichen Befunde werden spezifische Facetten der Gestaltung kontextpersonalisierter Aufgaben diskutiert. Während bei einer oberflächlichen Kontextpersonalisierung einzelne Oberflächenmerkmale, wie z.B. Namen der Protagonisten oder Orte, personalisiert werden, integriert eine tiefergehende Kontextpersonalisierung Interessensbereiche der Zielgruppe in die Aufgabenstellung, die für den Lernprozess sinnvoll genutzt werden können. Walkington und Hayata (2017) argumentieren, dass eine stark angereicherte Kontextpersonalisierung zu ablenkenden Effekten im Sinne des *seductive details effect* (Harp & Mayer, 1998) führen kann. Eine anhand von spezifischen Instruktionsprinzipien optimal angereicherte Kontextpersonalisierung begünstige hingegen kognitive Erleichterungseffekte (siehe *Contextual Grounding*, Son & Goldstone, 2009) welche wiederum bei einer oberflächlichen Kontextpersonalisierung ausblieben.

Der Forschungsstand zu Effekten der interessensbasierten Kontextpersonalisierung basiert vorrangig auf quantitativen Studien in der Sekundarstufe. In diesen Studien werden die zugrundeliegenden kognitiven Prozesse nur indirekt durch Schüler*innenleistungen in standardisierten Tests erfasst und erklärt (Fancsali & Ritter, 2014; Ku & Sullivan, 2002; Walkington, 2013). Durch diese ergebnisorientierte Betrachtung werden die tatsächlichen Denkprozesse während der Aufgabenbearbeitung nicht abgebildet. In einer qualitativ angelegten Studie sollen Denkprozesse von Grundschulkindern bei der Bearbeitung unterschiedlich stark kontextualisierter Aufgaben rekonstruiert werden. Ziel ist es, Erkenntnisse zu ablenkenden oder lern- und motivationsförderlichen Aspekten von kontextpersonalisierten Aufgaben aus Grundschüler*innenperspektive zu generieren.

Fragestellung

Welche Unterschiede beschreiben Kinder hinsichtlich ihrer kognitiven und motivationalen Prozesse bei der Bearbeitung unterschiedlich stark kontextualisierter Lernaufgaben?

Methodik

Zur Überprüfung der Fragestellungen ziehen wir als exemplarischen Lerngegenstand die Variablenkontrollstrategie (Chen & Klahr, 1999) heran, welche einen typischen Inhalt des frühen naturwissenschaftlichen Sachunterrichts darstellt. Die Datenerhebung findet in Einzeltreffen mit $N = 6$ Grundschulkinder der 4. Klasse statt. Nach einer kurzen Input-Phase zur Variablenkontrollstrategie werden zwei Aufgabenblätter zum Lerngegenstand dargeboten, welche hinsichtlich der Kontextpersonalisierung unterschiedlich gestaltet sind. Die Kontexte des stark interessensbasierten Aufgabenblattes orientieren sich an den Ergebnissen einer Erhebung aktueller Schüler*inneninteressen (z.B.: Fußballspielen, Haustiere; siehe: Laufs & Kempert, 2021b). Das oberflächlich kontextpersonalisierte Aufgabenblatt enthält keine expliziten Interessensbezüge (angelehnt an: Flugzeug-Aufgabe, siehe Bullock & Ziegler, 1999). Beide Aufgabenblätter werden unter Verwendung der Methode des Lauten Denkens (Konrad, 2020) bearbeitet, indem die Kinder gebeten werden, alles laut auszusprechen, was sie bei der Bearbeitung der Aufgabe denken. Es schließt ein Leitfadeninterview an, indem Nachfragen zum Bearbeitungsprozess der Aufgabenblätter

sowie zu motivationalen Aspekte bei der Auswahl und Bearbeitung der Aufgabenblätter thematisiert werden.

Analysen

Die Methode des Lauten Denkens (Konrad, 2020) soll offenlegen, welche Denkprozesse bei der Bearbeitung der Aufgaben ablaufen. Sowohl die Protokolle des Lauten Denkens als auch die Interviewtranskripte werden mit Hilfe der inhaltlich strukturierenden qualitativen Inhaltsanalyse nach Kuckartz (2016) u.a. anhand der deduktiven Kategorien *Erkennen der Interessensbezüge, Ablenkung durch Aufgabengestaltung* und *Motivation zur Auseinandersetzung mit dem Lerngegenstand* analysiert. Zudem werden induktive Kategorien abgeleitet.

Ausblick zur Ergebnispräsentation

In dem Beitrag soll der theoretische und empirische Hintergrund zu kognitiven und motivationalen Prozessen einer interessensbasierten Kontextpersonalisierung dargestellt werden. Ein Fokus liegt auf der Präsentation und Diskussion der Gestaltung der verwendeten Aufgabenblätter. Zudem sollen erste Ergebnisse aus der Pilotierung mit Bezug zur Forschungsfrage dargestellt und Schlussfolgerungen für die Hauptstudie gezogen werden. Aus den Ergebnissen sollen perspektivisch Gestaltungshinweise für eine lern- und motivationsförderliche interessensbasierte Kontextpersonalisierung abgeleitet werden.

5. Sprachbewusstheitsförderung beim mathematischen Modellieren

Daniela Balk
Universität Regensburg, Deutschland

Ziel

In diesem Posterbeitrag wird das Forschungsvorhaben *mamola* (*ma*thematical *mo*delling and *l*anguage *a*wareness) vorgestellt. Als wesentliches Erkenntnisinteresse dient die Frage, wie sich ein Arbeitssetting mit unterschiedlich hohem Grad an Sprachbewusstheitsförderung auf die mathematische Modellierungskompetenz von Grundschulkindern auswirkt.

Theoretischer Hintergrund

In der Grundschule ist mathematisches Modellieren bereits in allen Jahrgangsstufen möglich (Kolter und Eilerts 2014) und schafft eine wichtige Basis für komplexeres Modellieren an weiterführenden Schulen (van der Velden und Eilerts 2019). Allerdings zeigen die Kinder Schwächen in den Schritten des Modellierungskreislaufs, vor allem bei den Übergängen zwischen realer und mathematischer Welt (Eilerts und Kolter 2015; van der Velden und Eilerts 2019). Da häufig Kinder mit schwachen Sprachkompetenzen auch schwache Modellierer*innen sind, sprechen sich Specht und Tokarski (2019) in Bezug auf die Verbesserung der mathematischen Modellierungskompetenz für Sprachfördermaßnahmen aus. Diese konkretisieren sich bei Dröse (2019) in einer Sprachbewusstheitsförderung durch Formulierungsvariation, die den Modellierungsprozess im Bemerken wesentlicher Strukturen in den Textaufgaben nachweislich unterstützt. Als weitere empirisch belegte Fördermaßnahmen für Sprachbewusstheit eignen sich neben der Formulierungsvariation (Dröse 2019) Scaffolding (Jung und Chang 2007), korrektives Feedback

(Kavaliauskienė et al. 2007; Sato und Ballinger 2012) und die Selbstreflexion (Puengpipattrakul 2013).

Fragestellung/Erkenntnisinteresse

Für die vorliegende Studie sind drei Fragestellungen zentral. Als erstes Forschungsinteresse gilt, inwiefern in ihrer Sprachbewusstheit geförderte Kinder Modellierungsaufgaben besser lösen als vor der Intervention. Weiterhin ist herauszufinden, ob ein Setting mit einem hohen Grad an Sprachbewusstheitsförderung die Leistungen von schwachen Modellierer*innen mehr verbessert als ein niedriger Grad. Zuletzt wird untersucht, inwieweit sich die Kinder mit schwacher Modellierungskompetenz aus beiden Experimentalgruppen in ihrer Sprachbewusstheit unterscheiden.

Methodischer Zugang

Die Studie ist im Mixed-Methods-Design angelegt. Die ersten beiden Fragen werden quantitativ zu drei Messzeitpunkten untersucht. Aus den Ergebnissen der ersten Testung werden die schwachen Modellierer*innen selektiert, um deren Sprachbewusstheit für die dritte Forschungsfrage qualitativ mit selbst konzipierten, sprachbewussten Interviews zu erheben.

(Erwartete) Ergebnisse

Die Intervention sowie die Vor- und Nachtests fanden von Mai bis Juli 2022 statt. Die Follow-up-Tests folgen im Oktober/November 2022, so dass zum Zeitpunkt der Tagung bereits inferenzstatistische Ergebnisse präsentiert werden können, die Aussagen über die Effektivität von Fördermaßnahmen der Sprachbewusstheit auf die mathematische Modellierungskompetenz von Grundschulkindern möglich machen.

6. Förderung protokollbezogener Schreibfertigkeiten im Kontext von scientific inquiry

Jan-Martin Österlein[1], Mathias Ropohl[1], Sebastian Habig[2], Miriam Morek[1]
[1]Universität Duisburg-Essen, Germany; [2]Friedrich-Alexander Universität Erlangen-Nürnberg, Germany

Theoretischer Hintergrund

Textproduktion trägt im naturwissenschaftlichen Unterricht zum Erwerb einer *scientific literacy* bei (Wellington & Osborne, 2001). Mit *Writing to Learn* und *Learning to Write* werden zwei konträre Ansätze zur Integration von Schreibprozessen in den (naturwissenschaftlichen) Unterricht unterschieden. Während *Writing to Learn* auf reflexive Schreibaufgaben mit dem Ziel einer intensiven Auseinandersetzung mit naturwissenschaftlichen Konzepten zurückgreift, stehen bei *Learning to Write* die Merkmale naturwissenschaftlicher Fachsprache und fachtypischer Textsorten mit dem Ziel im Mittelpunkt, disziplinspezifische Schreibfertigkeiten wie z. B. argumentatives Schreiben, zu fördern (Sampson et al., 2013; Schmölzer-Eibinger & Thürmann, 2015). Zu diesem Zweck nutzt dieser Ansatz traditionelle Textsorten wie das Versuchsprotokoll.

Zielsetzung und Forschungsfragen

Vor dem beschriebenen Hintergrund untersucht das beschriebene Promotionsprojekt, inwiefern die Kombination von epistemologischen (*Writing to Learn*) und sprachlichen (*Learning to Write*) Schreibfördermaßnahmen die protokollbezogenen Schreibfertigkeiten im Chemieunterricht und den

Lernzuwachs hinsichtlich des chemischen Fachwissens fördert. Folgende Forschungsfrage soll im Rahmen des Projekts beantwortet werden:

Inwiefern beeinflusst die Kombination aus epistemologischen und sprachbezogenen Schreibfördermaßnahmen die protokollbezogenen Schreibfertigkeiten von Lernenden im Chemieunterricht?

Methode

Als Schreibfördermaßnahmen werden Kombinationen aus epistemologischen und fachsprachenbezogenen Fördermaßnahmen in einer experimentellen Interventions-studie in einem Prä-Post-Testdesign untersucht. Zwei von Leisen (2010) vorgeschlagene Methoden dienen dabei als Grundlage für sprachliche Hilfen:

1. Schreiben mithilfe von **Wortlisten** und **Satzbausteinen**
2. Schreiben anhand eines **Beispieltextes**

Die Kombination der epistemologischen und sprachlichen Fördermaßnahmen ergibt unter Berücksichtigung einer Kontrollgruppe ein 2x3-Gruppen-Design. Die erste Erhebungsphase umfasst n = 109 vollständige Datensätze von Lernenden der neunten Jahrgangsstufe. Weitere n = 300 Lernende werden in der zweiten Erhebungsphase (Herbst 2022) rekrutiert.

Im Prä- und Post-Test werden die Qualität der Versuchsprotokolle und das Fachwissen als abhängige Variablen erhoben, zusätzlich werden allgemeine Sprachfähigkeiten und figurale kognitive Fähigkeiten nach Heller und Perleth (2000) als Kontrollvariablen erhoben. Zu jedem Interventionszeitpunkt fertigen die Lernenden Versuchsprotokolle auf Basis von zwei Experimental-Videos an.

Ergebnisse und ihre Bedeutung

Zur Charakterisierung von Versuchsprotokollen ist ein Kodiermanual aus der Literatur abgeleitet und induktiv im Rahmen einer Vorstudie erweitert worden. Während der erste Teil des Kodiermanuals rein epistemologische Kategorien (Vollständigkeit, Überprüfbarkeit der Hypothese) umfasst, enthält der zweite Teil epistemologische Kategorien, die unmittelbar fachsprachlich umgesetzt werden. Das epistemologische Kriterium „akkurate Handlungsbeschreibung“ wird fachsprachlich etwa durch die Nutzung experimentbezogener Verben realisiert. Hinsichtlich der Interraterreliabilität konnte mit Werten für Cohen’s κ zwischen κ = .61 und .80 eine substanzielle Übereinstimmung erreicht werden (Landis & Koch, 1977).

Die protokollbezogenen Schreibfertigkeiten werden in einer ersten Analyse durch die Summenscores aller rein epistemischen Kategorien einerseits sowie aller epistemisch-sprachlichen Kategorien andererseits operationalisiert. Zur Untersuchung des Einflusses der Interventionsgruppe wurde eine ANOVA mit Messwiederholung berechnet. Die Ergebnisse zeigen keinen signifikanten Einfluss der Interventionsgruppe auf die Veränderung des Summenscores der epistemischen Kategorien, $F(5, 103) = .746$, $p = .591$. Auch für die Veränderung der epistemisch-sprachlichen Kategorien kann insgesamt kein signifikanter Einfluss festgestellt werden, $F(5, 103) = .084$, $p = .995$.

Auf Ebene des Protokollabschnitts können jedoch Unterschiede festgestellt werden. Für die epistemischen Kategorien der Beobachtung zeigt sich ein marginal signifikanter Unterschied

zwischen den Gruppen, $F(5, 103) = 2.11$, $p = .070$. Die sprachliche Qualität der Fragestellung nimmt über alle Gruppen hinweg signifikant zu, $F(1, 103) = 5.44$, $p = .022$, einen Unterschied zwischen den Gruppen gibt es jedoch nicht. Die sprachliche Qualität der Beobachtung, $F(1, 103) = 5.01$, $p = .027$, sowie der Auswertung, $F(1, 103) = 4.17$, $p = .044$, nimmt signifikant über alle Gruppen hinweg ab.

Aufgrund der noch geringen Stichprobe sind ausbleibende signifikante Effekte zwischen den Gruppen nicht überraschend. Jedoch entsprechen signifikante Verschlechterungen in den berichteten Protokollabschnitten nicht den Erwartungen. Durch eine qualitative Analyse des bearbeiteten Fördermaterials soll zudem das Nutzungsverhalten der Proband:innen untersucht werden, um zu überprüfen, inwieweit die berichteten Ergebnisse tatsächlich auf das Fördermaterial zurückzuführen sind.

7. Die Entwicklung einer Bildungsdokumentation zur Erfassung früher naturwissenschaftlicher Potenziale im Übergang Kita – Grundschule

Freya Müller, Hilde Köster
Freie Universität Berlin, Deutschland

Bereits im Elementarbereich können sich domänenspezifische Begabungspotenziale bei Kindern zeigen (Fuchs, 2019; Müller et al., 2021). Institutionsübergreifende Bildungsdokumentationen können im Übergang zur Grundschule eine Schlüsselfunktion in Hinblick auf eine kontinuierliche Förderung dieser Potenziale einnehmen (Fuchs, 2018; Holling et al., 2015; Köster et al., 2022). Bildungspolitische Bestimmungen stellen zudem die Bedeutsamkeit des Erfassens und der Förderung früher Begabungspotenziale in pädagogischen Kontexten heraus (Holling et al., 2015). Bisher eingesetzte Dokumentationsinstrumente für den Übergangsprozess betreffen vornehmlich mathematische und sprachliche Kompetenzbereiche (Fuchs, 2019; Geiling et al., 2015; Lingenauber & Niebelschütz, 2010). Auch hinsichtlich der Forschungsperspektive auf frühe Begabungspotenziale im Übergangsprozess kann derzeit noch eine mathematische Schwerpunktsetzung festgestellt werden (Bugzel, 2017; Fuchs, 2018; Heißenberger et al., 2017). Zunehmend wird jedoch auch die Relevanz des naturwissenschaftlichen Bildungsbereichs erkannt (BMBF, 2019; Köster, 2007; Steffensky, 2017). Wie naturwissenschaftliche Potenziale allerdings institutionsübergreifend dokumentiert und gefördert werden können, bleibt unklar.

Im Rahmen eines Teilprojekts des Forschungsverbunds „Leistung macht Schule (LemaS)“ wird der Fragestellung nach Anforderungen für eine praxistaugliche Bildungsdokumentation zur Potenzialerfassung im Übergangsprozess nachgegangen. Um den Anspruch einer anwendungs- und theoriebasierten Bildungsdokumentation zu gewährleisten, wurde ein iterativer Design-Based Research (DBR) Ansatz gewählt (Reinmann, 2017). Unter der leitenden Forschungsfrage:

Wie muss eine Bildungsdokumentation zur pädagogischen Beobachtung von naturwissenschaftlichen (Leistungs-)Potenzialen für den Übergang von der Kita in die Grundschule konzipiert sein, damit (potenziell-)begabte Kinder von pädagogischen Fachkräften prozessorientiert erkannt werden können?

wird das Forschungsvorhaben in drei methodisch explorativen und qualitativen Teilstudien realisiert. Dabei legt die Teilstudie 1 einen Fokus auf die Analyse der Bedarfe der Akteur*innen in den teilnehmenden Einrichtungen hinsichtlich der Übergangsgestaltung sowie der naturwissenschaftlichen

Potenzialerfassung. In dieser Projektphase wurde ein erster theoriegeleiteter Prototyp einer Bildungsdokumentation zur Potenzialerkennung entwickelt. In der Teilstudie 2 wird mithilfe formativer Evaluation gemäß DBR die Bildungsdokumentation iterativ weiterentwickelt und finalisiert. Die Teilstudie 3 verfolgt die Überprüfung der Nützlichkeit der entwickelten Bildungsdokumentation (Moosbrugger & Kelava, 2020).

Zur Realisierung der Teilstudien werden Daten aus vier vorschulischen Einrichtungen und drei Grundschulen von pädagogischen Fach- und Lehrkräften unterschiedlicher Bundesländer (Berlin, Hamburg, Schleswig-Holstein, Niedersachsen) herangezogen. Das methodische Vorgehen des Projekts gliedert sich in aufeinander folgende Phasen und Zyklen auf (DBR). Diese werden durch verschiedene qualitative Erhebungs- und Auswertungsmethoden empirisch begleitet.

Durch die Angaben aus der Teilstudie 1 zu Beginn des DBR-Zyklus mithilfe leitfadengestützter Interviews (*N* = 12 Pädagog:innen) und deren Auswertung mittels qualitativer Inhaltsanalyse nach Kuckartz und Rädiker (2020) konnte einerseits die Relevanz des Forschungsvorhabens zur Konzipierung einer Bildungsdokumentation zur frühen Potenzialerkennung gestützt werden. So zeigte sich, dass teilweise keine Dokumentation für den naturwissenschaftlichen Bereich genutzt wird. Überdies gab es in der Praxis auch keine explizite Schwerpunktsetzung hinsichtlich einer naturwissenschaftlichen Potenzialerfassung. Zudem gaben die Pädagog:innen wichtige Hinweise hinsichtlich der Konzipierung eines nützlichen Instruments für die Praxis. So ließen diese erkennen, dass geschlossen-strukturierte sowie auch offene Elemente für eine Anwendung wichtig sind. Auf Grundlage der Angaben wurde das Instrument zur Potenzialerfassung (weiter-)entwickelt und in der Praxis erneut erprobt. Die erste Sichtung der Daten aus Teilstudie 2 geben Hinweise darauf, dass durch die Einführung des Instruments eine Sensibilisierung hinsichtlich des naturwissenschaftlichen Bildungsbereichs bei den pädagogischen Fachkräften stattfindet. Überdies schätzen die Akteur*innen die entwickelte Bildungsdokumentation als hilfreich dahingehend ein, naturwissenschaftliche Potenziale im pädagogischen Alltag erkennen zu können.

Das Poster wird die qualitativen Ergebnisse der Teilstudie 1 sowie erste Ergebnisse der Teilstudie 2 vorstellen.

8. Frühe naturwissenschaftliche Bildung: Vignettentest

Eva Bühler[1], Bettina Grab[1], Markus Rehm[1], Hendrik Lohse-Bossenz[2], Tim Billion-Kramer[3]
[1]PH Heidelberg; [2]Universität Greifswald; [3]PH Ludwigsburg

Theoretischer Hintergrund: Die Förderung einer naturwissenschaftlichen Bildung soll bereits im frühen Kindesalter beginnen. Dies wird in den letzten Jahren vermehrt unter dem Begriff *frühe naturwissenschaftliche Bildung* (fnB) gefordert. Zur Förderung einer fnB können naturwissenschaftliche Denk- und Arbeitsweisen (nDuA) bereits in Kindertagesstätten ins Spiel kommen (Leuchter, 2017). Ziel ist es, naturwissenschaftliches Interesse und kindliche Neugier zu kultivieren, um sich die Welt *auch* naturwissenschaftlich erschließen zu können (Steffensky, 2017). Diese Forderungen nach einer fnB stellen anspruchsvolle berufliche Anforderungen an pädagogische Fachkräfte. Pädagogische Fachkräfte benötigen u.a. fachwissenschaftliche und fachdidaktische Kompetenzen, damit sie geeignete Lerngelegenheiten im Kindergarten anbieten können (Steffensky

et al., 2012). Wie pädagogische Fachkräfte Professionswissen zu nDuA im Studium sowie in Aus- und Fortbildung wirksam aufbauen, ist bislang wenig untersucht. Dies ist u.a. auf einen Mangel an (handlungsnahen) Erhebungsinstrumenten zurückzuführen, die ein solches Professionswissen valide erfassen können. In der Lehrkräftebildung haben sich bei der Erfassung professioneller Wissensinhalte sogenannte Vignettentests in verschiedenen Formaten (Text, Comic und Video) als unterschiedlich geeignet erwiesen (vgl. Brovelli et al., 2014; Rutsch, 2016; Friesen, 2017; Syring et al., 2015). Befunde zu unterschiedlichen Vignettenformaten zeigen, dass die Einschätzung der wahrgenommenen Authentizität, Motivation, Immersion und Resonanz der Vignettenformate durch die Probad:innen ein geeignetes Maß darstellt, die Qualität des handlungsnahen Erhebungsinstruments einschätzen zu können (Friesen, 2017; Friesen, Kuntze & Vogel, 2018; Seidel et al., 2011). Denn Authentizität, Immersion, Motivation und Resonanz sind wichtige Indikatoren für die kognitive Auseinandersetzung mit den vorgelegten Unterrichtssituationen (Friesen, Kuntze & Vogel, 2018). Immersion ist dabei ein Maß für das Hineinversetzen in die in der Vignette vorgelegte Unterrichtssituation, Resonanz spiegelt die Verbindung der Unterrichtssituation mit eigenen Lehr-Lernerfahrungen wider (Seidel et al., 2011; Friesen, Kuntze & Vogel, 2018). Für den frühpädagogischen Bereich stehen Untersuchungen der Qualität von Handlungssituationen in Vignetten anhand der Indikatoren Authentizität, Immersion, Motivation und Resonanz noch aus.

Fragestellung und Methode: Ziel des Projekts *Explore Scientific Inquiry– Kindergarten* (ESci-K) ist die Entwicklung eines Vignettentests zu folgenden nDuA: Fragen stellen, Vermuten, Beobachten, Ordnen und Systematisieren, Messen, Untersuchungen planen und durchführen. Pädagogische Fachkräfte werden gebeten, gezeigte Handlungsalternativen auf ihre Geeignetheit hin einzuschätzen fnB zu fördern. Eine Musterlösung wird theoriegestützt erstellt und zur Validierung an Expert:innen von fnB und Naturwissenschaftsdidaktiken weitergeleitet. Weiterhin werden je Format Items zu Authentizität, Immersion, Motivation und Resonanz erhoben, um so Aussagen zu Wahrnehmung und Geeignetheit der verschiedenen Formate bei Aus-, Fort- und Weiterbildung von pädagogischen Fachkräften treffen zu können. Folgende Fragstellung wird fokussiert: Wie werden unterschiedliche Vignettenformate (Text, Comic, Video) von den pädagogischen Fachkräften in Bezug auf Authentizität, Immersion, Motivation und Resonanz wahrgenommen?

Erwartete Ergebnisse: Es wird erwartet, dass die Formate Text, Video und Comic unterschiedlich bewertet werden. Aus bisherigen Vergleichsstudien lassen sich für Textvignetten z.B. Vorteile in der Zeitlichkeit, der niedrigen Individualität und des Ausbleibens von Störfaktoren festmachen. Gleichzeitig kann die reduzierte Menge an Kontextinformation in der Wahrnehmung zu weniger Authentizität und Immersion führen. Bei den Videovignetten gestaltet sich die Zeitlichkeit ähnlich dem Kita-Alltag, obgleich ein wiederholtes Abspielen möglich ist. Sie zeigen Kinder und pädagogische Fachkräfte in alltagsnahen Situationen, dadurch werden höhere Werte in den Konstrukten Authentizität, Immersion und Motivation erwartet. Andererseits stellen die Videovignetten höhere Anforderungen an die pädagogischen Fachkräfte. Sie müssen die relevanten Informationen aus dem bewegten Bild extrahieren, ohne durch die große Menge an Kontextinformationen abgelenkt zu werden. Die Comicvignetten werden als Format „dazwischen“ gesehen, sie sind in Bezug auf Zeitlichkeit mit Texten vergleichbar, können aber auch in Bezug auf Individualität mit Videos

verglichen werden. Die Darstellungsform erlaubt es, Mimik zu erkennen, reduziert jedoch die Individualität der Figuren, was die Analyse aufgrund der geringeren Kontextinformationen erleichtern könnte (vgl. Friesen 2017; Friesen, Kuntze & Vogel 2018). Die zu erwartenden Ergebnisse liefern wichtige Hinweise über die Geeignetheit verschiedener Vignettenformate im frühpädagogischen Bereich.

9. Geschichte lernen digital – Wie lösen Schüler*innen historische Lernaufgaben der Plattform segu?

Lena Liebern
Universität Duisburg-Essen, Deutschland

Im Laufe der letzten Jahre kann die Disziplin der Geschichtsdidaktik auf eine Breite an Beiträgen zum digitalen historischen Lernen zurückblicken, die jedoch nach wie vor weitgehend theoretisch geprägt sind (Bernsen, König & Spahn, 2012; Demantowsky & Pallaske, 2015). Im Bereich der empirischen Forschung dominieren Beiträge zu Einzelphänomenen mit Schwerpunkt auf geschichtskulturellen Aspekten, wie z. B. dem Umgang mit der Wikipedia (Hodel, 2013) oder Social Media (Burkhardt, 2021). Hingegen sind empirische Forschungen spezifisch zu historischen Lernmedien bzw. zu der Frage nach dem schulischen historischen Lernen in einer zunehmend digitalisierten Welt nach wie vor unterrepräsentiert (Degner & Schäfer, 2010). Schwabe (2018: 422) konstatierte, dass die Überlegungen zum digitalen historischen Lernen eine empirische Unterfütterung benötigen, „um geschichtsdidaktisch begründete digitale Unterrichtsmaterialien und Lehr-Lern-Arrangements konzipieren zu können." Erst ein genaueres Wissen über Lernvorgänge an, mit und in digitalen Medien erlauben Auskunft über die potenziellen Chancen des digitalen historischen Lernens.

Dieses Forschungsdesiderat wird durch das Projekt „Geschichte lernen digital" angegangen: Durch eine *Mixed-Method-Studie* mit der Variante eines *explanatory designs* (Prinz & Thünemann, 2016) wird zum einen das Lernangebot der Lernplattform *segu* durch eine quantitative Analyse ausgewertet, indem die Lernaufgaben (n = 339) und Lernmaterialien typologisiert werden. Daraus ergibt sich die Forschungsfrage, welchem Lernbegriff und welcher Aufgabenkultur das webbasierte Angebot *segu* folgt. Dabei spiegelt die analytische Auswertung das Bild anderer Lernaufgabenanalysen wider (Michler, 2007; Waldis, Hodel & Fink, 2012): Über 50% der analysierten Lernaufgaben beziehen sich auf textförmige vorgegebene Materialien, sind definiert und konvergent konstruiert und zielen auf den Erwerb reproduktiver, faktenbasierter Wissenselemente und prä-narrativer Teilkompetenz ab. Die Analyse verweist darauf, dass auch in digital geprägten Lernumgebungen eine Textlastigkeit der Materialien vorliegt und interaktive Aufgabenformate (z. B. Quizformate) in ihrem Nutzen für das historische Lernen begrenzt sind.

Basierend auf dem Analyseergebnis der quantitativen Studie wurden für die qualitative Untersuchung drei Lerneinheiten ausgewählt, in denen eine Graduierung der Aufgabenniveaustufen erkennbar ist und verschiedene digitale Tools, bspw. eine interaktive Bildquelle oder eine *Street-View*-Anwendung, implementiert sind. Ziel der qualitativen Studie im explorativen Design ist es (Mayring, 2010), Hypothesen über die Performanz historischer (Denk-)Operationen von Schüler*innen-Tandems zu generieren: Welche Praktiken der Aufgabenbearbeitung lassen sich im Lösungsprozess der Schüler*innen im Umgang mit verschiedenen digitalen Aufgabenformaten rekonstruieren? Dazu

wurden Schüler*innen-Tandems (*n* = 10) einer 7. und 8. Jahrgangsstufe zweier Gymnasien in Nordrhein-Westfalen bei der Bearbeitung eines Lernmoduls der Lernplattform *segu* videografiegestützt beobachtet (Wagner-Willi, 2013; Waldis, 2016). Als Auswertungsmethode wird auf die Dokumentarische Methode zurückgegriffen, da sie mit der Ausrichtung auf Prozesse der Wissenskonstruktion und -genese eine methodologisch fundierte Analyseeinstellung auf das Lösen von und Lernen mit digitalen historischen Aufgabenformaten bietet (Asbrand & Martens, 2018). Die qualitativ-rekonstruktive Erforschung zeigt, dass Praktiken der digitalen Aufgabenbearbeitung je nach Aufgabenmodus und Tandem äußerst heterogen und vielschichtig sind. Bislang konnten die *Tertia Comparationis* Metakommunikation über die Aufgabenstellung, Umgang mit (digitalen) Materialien, Operationen historischen Denkens und Digitale Strategien ermittelt werden. Im folgenden Schritt der sinngenetischen Typenbildung werden die Ergebnisse weiter verdichtet (Bohnsack, 2013).

10. Verstehensschwierigkeiten beim Lernen mit einzelnen Repräsentationen

Kerstin Gresens, Hendrik Härtig
Universität Duisburg-Essen, Deutschland

Unter Repräsentationen werden kognitive Werkzeuge zum Verstehen von Konzepten und dem Umgang mit diesen verstanden (Ainsworth, 2008). Es gibt verschiedene Repräsentationsformen wie Bilder, Diagramme oder Tabellen, die einen Sachverhalt beschreiben (Klein et al., 2018). In den Naturwissenschaften werden Repräsentationen eingesetzt, um abstrakte Konzepte zu vermitteln, Probleme zu verstehen, zu erkunden und zu lösen (Corradi et al., 2012). Für das Lernen ist es notwendig, die verschiedenen Repräsentationsformen interpretieren zu können und in Verbindung mit der Realität und dem zu lernenden Konzept zu bringen (Cock, 2012). Obwohl für viele Lernende die Verwendung von multiplen Repräsentationen schwierig ist (Ainsworth, 1999), werden diese häufig im Unterricht eingesetzt. Es ist sinnvoll, zunächst jede einzelne Repräsentationsform zu verstehen (Ainsworth, 2006). Inwieweit sich die Befunde aus vorherigen Studien auf universitärem Niveau (Nguyen & Rebello, 2009, 2011) oder in anderen Disziplinen (Dittmar et al., 2017) auf den Physikunterricht übertragen lassen, ist unklar, weswegen in der vorliegenden Studie Verstehensschwierigkeiten zu einzelnen Repräsentationsformen im Physikunterricht der Sekundarstufe 1 erhoben werden sollen. Zusätzlich ist bekannt, dass Lernende häufig keine Regelmäßigkeiten und Unterschiede zwischen verschiedenen Repräsentationsformen erkennen (Ainsworth, 2006). Dies könnte zu einer fehlerhaften Wahrnehmung und damit verbunden zu Verstehensschwierigkeiten führen. Dies resultiert in den folgenden Forschungsfragen:

FF1: Welche Verstehensschwierigkeiten ergeben sich bei der Nutzung einzelner Repräsentationen?

FF2: Wie nehmen Schülerinnen und Schüler verschiedene Repräsentationsformen wahr?

FF3: Inwieweit hängen Wahrnehmung und Verstehensschwierigkeiten zusammen?

Zur Beantwortung der Forschungsfragen wird eine qualitative Befragung mit den Methoden Lautes Denken (Deffner, 1984) und Repertory Grid Technik (Björklund, 2008) durchgeführt. Zur Beantwortung der ersten Forschungsfrage wird das Laute Denken beim Bearbeiten von Aufgaben, die das Verstehen einer Repräsentationsform erfordern, genutzt. Dazu bearbeiten die Lernenden Aufgaben mit einer bestimmten Repräsentation. Die während der Bearbeitung von Lernenden

erstellten Notizen werden mit den dazugehörigen verbalen Äußerungen mit einem Smartpen aufgenommen. Aus den transkribierten Äußerungen und Notizen schließen wir auf Verstehensschwierigkeiten. Diese werden dann induktiv nach Mayring erstellten Kategorien zugeordnet. Über die Repertory Grid Technik wird in der vorliegenden Studie die persönliche Wahrnehmung der einzelnen Repräsentationsformen für jeden Lernenden erfragt. Hierzu werden zunächst verschiedene Triaden von Elementen (je fünf Repräsentationsformen zu zwei verschiedenen Themen) von den Lernenden sortiert, indem zwei Elemente eine Gemeinsamkeit aufweisen, in der sich das dritte Element unterscheidet. So werden Konstrukte aufgestellt anhand derer im Anschluss alle Elemente auf einer fünfstufigen Skala bewertet werden (vgl. u. a. Scheer & Catina, 1993). Aufgrund des so entstandenen Grids können Aussagen über ein bestimmtes Element über alle Konstrukte bzw. vergleichende Aussagen über alle Elemente innerhalb eines Konstrukts getroffen werden. Die Erhebung findet im Oktober und November 2022 statt. Bis zur Präsentation des Posters auf der GEBF werden die Daten der Repertory Grid Technik ausgewertet sein und können präsentiert werden, Die Auswertung des Lauten Denkens ist vermutlich noch im Prozess, sodass erste Einblicke präsentiert werden können.

11. Räumliche Fähigkeiten in der Chemie – Mehr als nur mentale Rotation?

Sebastian Nickel, Steffen Brockmüller, Sebastian Habig
Friedrich-Alexander-Universität Erlangen-Nürnberg

In den naturwissenschaftlichen Domänen, besonders in der Chemie, spielen visuelle Repräsentationen eine Schlüsselrolle beim Verständnis von fachlichen Konzepten (Kozma & Russell, 2007) und beim Problemlösen (Rau, Zahn, Misback, Herder, & Burstyn, 2021). Repräsentationen werden u. a. genutzt, um Daten mithilfe von Diagrammen, Strukturen von Molekülen und Reaktionsmechanismen in 2D und 3D zu veranschaulichen (Harle & Towns, 2011). Hierzu werden Kompetenzen zum Umgang mit Repräsentationen benötigt, wie die Fähigkeit Repräsentationen ineinander zu überführen und Verbindungen zwischen ihnen zu erkennen (Kozma & Russell, 1997, 2007). Diese Repräsentationskompetenzen müssen in der Schullaufbahn und dem Studium entwickelt und gefördert werden. Räumliche Fähigkeiten spielen in diesem Zusammenhang eine wichtige Rolle und wirken sich positiv auf den Umgang und das Lernen mit Repräsentationen aus (Carter, LaRussa, & Bodner, 1987; Buckley, Seery, & Canty, 2018). Die Räumlichen Fähigkeiten bilden einen von 16 Faktoren ab, in die Intelligenz nach dem Cattell-Horn-Carroll-Modell untergliedert werden kann (Carroll, 1993; Schneider & McGrew, 2012). Diese lassen sich weiter in elf verschiedene räumliche Faktoren untergliedern (Schneider & McGrew, 2012), wie die mentale Rotation, die räumliche Visualisierung und die räumliche Orientierung (Buckley, Seery, & Canty, 2018). In den meisten Studien zum MINT-Bereich, in denen räumliche Fähigkeiten berücksichtigt werden, wird zumeist ein Faktor zugrunde gelegt, z. B. die mentale Rotation. Wie verschiedene Faktoren räumlicher Fähigkeiten die Bearbeitung repräsentationsbasierter Chemieaufgaben beeinflussen, ist aktuell jedoch noch wenig untersucht und bedarf weiterer Forschung. Daraus lässt sich folgende Forschungsfrage für den ersten Teil des Projekts ableiten:

Welche Faktoren räumlicher Fähigkeiten spielen beim Lösen repräsentationsbasierter Chemieaufgaben eine Rolle?

Zur Untersuchung dieser Forschungsfrage werden Test-Items im geschlossenen und halboffenen Aufgabenformat entwickelt und aus bereits bestehenden Testinstrumenten adaptiert (Averbeck, 2021; Taskin, Bernholt, & Parchmann, 2015). Die Test-Items decken verschiedene Teilgebiete der Chemie ab und werden anhand von sechs Repräsentationskompetenzen (adaptiert nach Kozma und Russell, 2007) und fachlicher Basiskonzepte der Chemie kategorisiert. Die Konstruktvalidität dieser Kategorisierung wird ermittelt, indem Fachkolleg:innen aus der Chemiedidaktik die Items mithilfe eines Manuals einordnen und die resultierenden Zuordnungen mithilfe der Interrater-Reliabilität auf Übereinstimmung geprüft werden. Im Wintersemester 2022/23 werden die Test-Items in Form eines Multiple Matrix Samplings (Childs & Jaciw, 2002) in Verbindung mit Testinstrumenten zur Erhebung der verschiedenen Faktoren räumlicher Fähigkeiten, u. a. Hidden Figure Test und Paper Folding Test (Ekstrom, French, Harman, & Derman, 1976), eingesetzt. Dadurch wird ein detaillierter Überblick über die Wirkmechanismen beim Bearbeiten von repräsentationsbasierten Chemieaufgaben erhalten. Die Studie wird mit (angehenden)
Chemiestudierenden durchgeführt. Der Zusammenhang zwischen den Faktoren räumlicher Fähigkeiten und spezifischen Repräsentationen und Repräsentationskompetenzen wird mithilfe von Korrelations- und Regressionsanalysen geprüft.

Basierend auf den Ergebnissen der ersten Forschungsfrage wird eine Interventionsstudie im Prä-Post-Design geplant. In dieser sollen die aus der ersten Forschungsfrage als relevant hervorgegangenen Faktoren räumlicher Fähigkeiten gezielt gefördert werden (vgl. Velázquez & Méndez, 2021) und geprüft werden, ob sich diese Förderung positiv auf den Lernerfolg auswirkt. Damit leisten die Ergebnisse dieses Projekts einen Beitrag zur Untersuchung der Wirkmechanismen von Faktoren räumlicher Fähigkeiten auf den Lernerfolg mit externen Repräsentationen in der Chemie. Des Weiteren können Möglichkeiten aufgedeckt werden, wie der Lernerfolg von Studierenden mit gezielter Förderung spezifischer Faktoren räumlicher Fähigkeiten verbessert werden kann.

12. Effekte der Covid-19 Pandemie auf die Mathematikleistung von Zweitklässlern in Deutschland

Ophelia Isabella Urbach, Boris Forthmann, Natalie Förster, Elmar Souvignier
Westfälische Wilhelms-Universität Münster, Germany

Theoretischer Hintergrund

In Reaktion auf die Covid-19 Pandemie und den damit verbundenen Wechsel in den Distanzunterricht wurden bereits zu Beginn der Pandemie Leistungseinbußen und eine höhere Leistungsheterogenität befürchtet (Leopoldina, 2020; Wößmann, 2020). Aktuelle Reviews und eine Meta-Analyse zeigen dabei überwiegend negative Auswirkungen der Pandemie auf kindliche Leistungen, wobei vor allem jüngere Kinder stärker betroffen sind (Betthäuser et al., 2022; Hammerstein et al., 2021; König & Frey, 2022). In zwei weiteren Reviews wurde zudem herausgestellt, dass vor allem die Mathematikleistung (im Gegensatz zur Leseleistung) stärker beeinträchtigt ist (Böttger & Zierer, 2021; Helm et al., 2021). Die Studien im deutschsprachigen Raum im Bereich Mathematik beschränken sich aktuell auf Kinder ab der vierten Klasse (Depping et al., 2021; Schult et al., 2022a, 2022b; Spitzer & Musslick, 2021; Tomasik et al., 2021), und es gibt bislang nur eine Studie in der die mittelfristigen

Auswirkungen im zweiten Schuljahr unter Pandemiebedingungen betrachtet wurden (Schult et al., 2022b). In unserer Studie untersuchen wir daher kurz- und mittelfristige Auswirkungen der Pandemie auf Kinder in der zweiten Klasse in Mathematik.

Fragestellung

Wie haben sich die durchschnittliche Mathematikleistung sowie interindividuelle Unterschiede in der Mathematikleistung in der zweiten Klasse durch die Corona-Pandemie verändert?

1. Die durchschnittliche Mathematikleistung ist nach Beginn der Pandemie geringer als vor der Pandemie.
2. Die interindividuellen Unterschiede haben sich während der Pandemie vergrößert.

Methode

Zur Beantwortung unserer Fragen vergleichen wir die mathematischen Leistungen von Zweitklässlern aus drei Teilstichproben: der ersten Pandemie-Kohorte (N = 3204) im Schuljahr 2019/20, der zweiten Pandemie-Kohorte (N = 2357) im Schuljahr 2020/21, sowie vier prä-Pandemie-Kohorten (N = 6294) aus den Schuljahren 2015/16 bis 2018/19. Die Daten wurden mit der web-basierten Lernverlaufsdiagnostik quop (Souvignier et al., 2021) erhoben. Die quop-M2 Testreihe besteht aus acht parallelen Tests, und erfasst die mathematische Vorläuferfähigkeit (vier Subfacetten) und Rechenfähigkeit (sechs Subfacetten; Salaschek & Souvignier, 2014). Diese Tests wurden in Abständen von drei Wochen über das Schuljahr verteilt bearbeitet. Die latenten Mittelwerts- und Varianzunterschiede wurden mit einem zweidimensionalen Modell mit den Faktoren mathematische Vorläuferfähigkeit und Rechenfähigkeit und den entsprechenden Subfacetten als Indikatoren mit dem R-Paket lavaan (Rosseel, 2012) geschätzt. Das Messmodell zeigte einen guten Modellfit (robust CFI = 0.966, robust RMSEA= 0.017) nach Hu und Bentler (1999) und eine strikte Messinvarianz über die Gruppen sowie eine partielle starke Messinvarianz (drei der sechs Intercepts für Rechenfähigkeit wurden frei geschätzt) über die Zeit nach Chen (2007).

Ergebnisse

Die Ergebnisse der durchschnittlichen Mathematikleistung fallen für die beiden Fähigkeitsbereiche unterschiedlich aus. In den Vorläuferfähigkeiten war die durchschnittliche Leistung in den beiden Pandemiekohorten zu fast allen Messzeitpunkten signifikant niedriger als vor der Pandemie. Allerdings zeigten sich die niedrigeren Vorläuferfähigkeiten bereits vor der Pandemie und die geringsten (und nicht signifikanten) Unterschiede während des ersten Lockdowns. Im Gegensatz dazu waren die arithmetischen Fähigkeiten in den beiden Pandemie-Kohorten im Vergleich zu den Prä-Pandemiekohorten teilweise signifikant höher, wobei die Unterschiede im Verlauf des Schuljahres abnahmen.

Die interindividuellen Unterschiede in den Vorläuferfähigkeiten waren in beiden Pandemiekohorten höher als zuvor, dies gilt jedoch ebenfalls für Messzeitpunkte vor Beginn der Pandemie. Bei den arithmetischen Fähigkeiten war die Leistungsheterogenität im Schuljahr 2019/20 vor Beginn der Pandemie signifikant höher und nahm während des ersten Lockdowns entgegen der Erwartung ab. Für die zweite Pandemiekohorte zeigte sich die erwartete Zunahme der interindividuellen Unterschiede in der Mathematikleistung.

Diskussion

Obwohl der Vergleich der beiden Pandemiekohorten keinen klaren negativen Effekt der Pandemie auf die durchschnittliche Mathematikleistung zeigt, deuten die höheren interindividuellen Leistungsunterschiede in beiden mathematischen Fähigkeitsbereichen trotzdem auf einen mittelfristigen negativen Einfluss der Corona-Pandemie hin. Da dieser Anstieg in den interindividuellen Leistungsunterschieden allerdings auch bereits vor der Pandemie zu finden war, ist es wichtig, die Pandemieauswirkungen in detaillierten längsschnittlichen Analysen differenziert für verschiedene Kompetenzbereiche und Altersstufen weiter zu untersuchen.

13. One fits all? Förderung selbstbestimmter Motivation bei einer Umweltbildungsmaßnahme im Wald

Melanie Basten[1], Daniela Sellmann-Risse[1], Silvia Fränkel[2]
[1]Universität Bielefeld, Germany; [2]Universität zu Köln, Germany

Theoretischer Hintergrund

Bildung für Nachhaltige Entwicklung (BNE) ist ein hochkomplexer Bildungsbereich mit der Zielsetzung, bei Schüler:innen Gestaltungskompetenz auszubilden (de Haan, 2008). Diese ermöglicht es, nachhaltige Handlungsentscheidungen zu treffen und umzusetzen (ebd.). Neben der Gestaltungskompetenz ist die selbstbestimmte Motivation im Sinne der Self-Determination Theory (SDT, Ryan & Deci, 2018) Prädiktor für nachhaltiges Handeln (Pelletier et al., 2011). Nicht zuletzt die Corona-Pandemie hat gezeigt, dass es wichtig ist, dass alle an naturwissenschaftlicher Grundbildung (Bennett et al., 2007) partizipieren, wenn es um gesellschaftlich relevante naturwissenschaftliche Themen (Sadler et al., 2007) geht (Basten & Großmann, 2022). Umweltbildungsmaßnahmen sollten dementsprechend auch alle Lernenden gleichermaßen erreichen. Es zeigte sich in einer vorangehenden Studie bereits, dass die hier untersuchte Umweltbildungsmaßnahme „Waldjugendspiele“ hinsichtlich Naturverbundenheit, einem weiteren Prädiktor nachhaltigen Handelns, bei Jungen und Mädchen gleich effektiv ist (Basten et al., 2021).

Fragestellung

In der aktuellen Studie soll untersucht werden, ob die Waldjugendspiele die Grundbedürfnisse erfüllen und so zu intrinsischem Erleben bei allen Kindern führen. Autonomie, Kompetenz und soziale Eingebundenheit sind dabei die universellen psychologischen Grundbedürfnisse aller Menschen, deren Erfüllung zu selbstbestimmten Motivationsformen, wie bspw. dem intrinsischen Erleben, führen kann (Ryan & Deci, 2018). Werden umweltbildungsbezogene Interaktionen mit der Natur wiederholt als grundbedürfnisunterstützend und selbstbestimmt erlebt, kann der Schutz der Natur dauerhaft in das Wertesystem der Lernenden aufgenommen werden (siehe Pelletier et al., 2011).

Methode

Stichprobe und Untersuchungsdesign

An der Umweltbildungsmaßnahme nahmen N=1585 Schüler:innen der 3. Klasse (9,2 Jahre; 50,7% Mädchen) teil. 56,2% der Kinder hatte einen Migrationshintergrund und ihre familiäre Herkunft ließ sich sieben Kulturarealen (Hunter & Whitten, 1976) zuordnen (siehe Fränkel et al., 2019). Die Schüler:innen besuchten im Klassenverband die Waldjugendspiele, die hauptsächlich auf die Naturerfahrung (Lude, 2021; siehe Sellmann-Risse et al., 2021) ausgerichtet sind und nur wenige

Bildungselemente enthalten, die Barrieren hinsichtlich Leseverständnis und Schriftsprache darstellen können (Fränkel et al., 2018).

Messinstrumente

Die Fragebögen wurden von den Schüler:innen am Ende des Waldganges mit der Hilfe von geschulten Lehramtsstudierenden ausgefüllt. Die wahrgenommene Erfüllung der Grundbedürfnisse sowie das intrinsische Erleben wurden mit der KIM (Kurzskala Intrinsische Motivation, Wilde et al., 2009; Skalen: wahrgenommene Kompetenz (WK), wahrgenommene Wahlfreiheit (WW) und Interesse/Vergnügen (IV)) und dem IMI (Intrinsic Motivation Inventory, Ryan, 1982; Skala: soziale Eingebundenheit (SE)) erhoben. Zusätzlich wurden die Natur- und Waldverbundenheit mit einer vereinfachten Inclusion of Nature in Self-Scale von Schultz (2002) erfasst, wobei für die Waldverbundenheit das Wort „Natur“ durch „Wald“ ersetzt wurde.

Datenanalyse

Mit einem Strukturgleichungsmodell wurde überprüft, ob die Waldjugendspiele motivierend für alle Kinder sind (RMSEA=.015; CFI=.985; TLI.=975; SRMR Within=.024; SRMR Between=.096). Die drei psychologischen Grundbedürfnisse sowie das intrinsische Erleben wurden als abhängige Variablen in das Modell eingefügt, Wald- und Naturverbundenheit sowie demografische Daten (Zugehörigkeit zu einer inklusiven Klasse, Geschlecht, kulturelle und regionale (Stadt/Land) Herkunft und Häufigkeit des Waldaufenthaltes) als Prädiktoren.

Ergebnisse und ihre Bedeutung

Insgesamt ist die Erfüllung der Grundbedürfnisse hoch (WW=2,6; Max.=4) bzw. sehr hoch (WK=3,4; SE=3,3) ausgeprägt. Auch das intrinsische Erleben der Kinder bei den Waldjugendspielen ist sehr stark (IV=3,5). Die Auswertung entlang des Strukturgleichungsmodells ergab signifikante Einflüsse verschiedener Prädiktoren auf die Skalen der KIM/IMI. Die Waldverbundenheit – und bezogen auf andere natürliche Umgebungen auch die Naturverbundenheit – sollte demnach bei allen Schüler:innen gestärkt werden, da sie einen Einfluss auf die Motivation bei einer Umweltbildungsmaßnahme hat (WW: .19**; WK: .39***; SE: .22*; IV: .34***). Eine Umweltbildungsmaßnahme, die besonders Naturerfahrung ermöglicht und wenige sprachliche Barrieren aufbaut, kann zu intrinsischer Motivation bei allen Schüler:innen führen, also im Prinzip passend für alle Kinder sein. Mädchen (SE: -.09*; IV: -.15***) und Kinder, die auch im Alltag eine höhere Nähe zur Natur haben (WK Landkinder: .13**; WW Waldbesuchshäufigkeit: -.11**), erleben die Waldjugendspiele dabei als motivierender, sodass auf Jungen und Kinder mit weniger Berührungspunkten mit der Natur besonderes Augenmerk gelegt werden sollte.

14. Referenzwortschatz für Schule: Sprachdidaktisch und sprachwissenschaftlich begründetes Wortschatz-Tool

Rebekka Wanka, Viktoria Michels, Michael Becker-Mrotzek
Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache, Germany

Das Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache der Universität zu Köln entwickelt derzeit einen *Referenzwortschatz für die Schule* für die Klassenstufen 1 bis 10. Die Datenbank wird in Form eines webbasierten Online-Tools voraussichtlich Ende 2022 veröffentlicht

werden und sie umfasst ungefähr 5.000 Wörter, mit jeweils ca. 120 begrifflich vernetzten Informationen je Eintrag. Diese vernetzten Informationen umfassen zahlreiche Parameter beispielsweise zu den Bereichen Aussprache, thematische Felder, Grammatik, Herkunft, Orthografie und Häufigkeit. Eine Vielzahl von Filtern innerhalb der Datenbank macht es möglich, Wortlisten aus drei Teilwortschätzen (1. Orthografiewortschatz, 2. Schreibwortschätze, 3. projekteigenes Schulbuchkorpus) für unterschiedliche Zwecke und Fächer in Schule, Bildung und Wissenschaft abzufragen.

Das interdisziplinäre Tool *Referenzwortschatz für die Schule* ist für unterschiedliche Nutzergruppen wie Lehrkräfte und Bildungseinrichtungen, Verlage, Wissenschaft und Lernende von Interesse: Lehrkräften können es für die individuelle Förderung von Schülerinnen und Schülern und die Unterrichtsplanung und -entwicklung nutzen. Verlage können die Datenbank für die Entwicklung von Schulbüchern und Wissenschaftlerinnen und Wissenschaftler für die Forschung sowie beispielsweise für die Entwicklung von Diagnostika verwenden. Zudem kann das Tool auch für Lernerinnen und Lernern – angeleitet durch Lehrkräfte – bei der Aneignung und Erweiterung eines bildungssprachlichen Registers hilfreich sein.

Lehrkräfte können den Aufbau und die Entwicklung eines bildungssprachlichen Wortschatzes unterstützen, indem sie diesen bei der Unterrichtsplanung und -gestaltung gezielt einsetzen. Bildungssprachliche Mittel betreffen sowohl den Wortschatz und syntaktische Strukturen als auch komplexe sprachliche Handlungen. Sie werden häufig für alle Schülerinnen und Schüler vorausgesetzt, aber nicht explizit in der Schule vermittelt (vgl. Feilke, 2012). Lernende aus sozioökonomisch privilegierten Familien und mit hohen bildungssprachlichen Kompetenzen entwickeln zumeist höhere Fachkompetenzen (de Moll 2018), was eine Benachteiligung von Schülerinnen und Schüler und Schülerinnen aus mehrsprachigen und/oder sozioökonomisch schlechter gestellten Familien bedeutet. Der Referenzwortschatz für die Schule kann eine Orientierung geben, welche Wörter im Schulkontext häufig und/oder relevant für einen bildungssprachlichen Wortschatz sind und damit einen Beitrag zur Bildungsgerechtigkeit leisten.

In unserem Poster stellen wird den *Referenzwortschatzes für die Schule* vor. Dabei wird neben der Kategorien- und Suchfilterentwicklung auf die nutzerorientierte Umsetzung und Gestaltung der Datenbank und Datenbankfilter eingegangen. Zudem werden mögliche didaktische Nutzungs- und Unterstützungsmöglichkeiten des digitalen Tools im Schulkontext für den Ausbau und die Entwicklung bildungssprachlicher Kompetenzen von Schülerinnen und Schüler vorgestellt.

15. Einfluss von Tiefe und Breite des Wortschatzes auf die Schreibflüssigkeit und die Textproduktion

Markus Linnemann[1], Sabine Stephany[2]
[1]Universität Koblenz, Germany; [2]Universität zu Köln, Germany

Die Produktion von Texten ist ein kognitiv anspruchsvoller Vorgang. Die deklarativen und prozeduralen Wissensbestände, die beim Schreiben relevant sind, sind außerordentlich vielfältig: Vom prozeduralisierten Wissen der Graphomotorik bis hin zur bewussten und zielgerichteten Antizipation des Adressaten sind verschiedene Komponenten beteiligt. Ein genauerer Blick auf Modelle des Schreibprozesses (z.B. Hayes, 2012) zeigt zudem, dass das Abrufen von Wörtern eng

mit der Textproduktion, insbesondere aber mit dem flüssigen Schreiben verbunden ist. Schreibflüssigkeit besteht zum einen aus der Übersetzung vorsprachlichen, propositionalen Gehaltes in linearisierte Sprache (Formulierungsflüssigkeit) und zum anderen aus dem schnellen Abruf und dem Verschriften von Buchstaben, Wörtern und Sätzen (Transkriptionsflüssigkeit) (Linnemann et al., 2022; McCutchen, 1996). Beide Prozesse sind in hohem Maße auf Wortschatzressourcen angewiesen. Es kann damit die Hypothese formuliert werden, dass die Textproduktion, insbesondere die Schreibflüssigkeit substanziell mit dem Wortwissen korrelieren.

In einer ersten korrelativen Studie wurde an einer Stichprobe von 129 Viert-, Sechst- und Neuntklässlern neben Variablen der Schreibflüssigkeit und der Textqualität als Maß für die allgemeine Schreibkompetenz der Peabody Picture Vocabulary Test PPVT (Dunn & Dunn, 2015) durchgeführt. Die Korrelation des Wortwissens mit der Schreibflüssigkeit ($r_{Transkriptionsflüssigkeit}$=.49; $r_{Formulierungsflüssigkeit}$=.29) und mit der Textqualität (r=.43) war wie angenommen ausreichend hoch. Trotz dieser moderaten Korrelationen trug das Wortwissen im Rahmen eines umfangreichen Strukturgleichungsmodells wenig zur Schreibflüssigkeit und zur Textproduktion bei. Einer der Gründe dafür ist möglicherweise eine fehlende Differenzierung des Konstrukts Wortschatz in die Facetten „Tiefe“ und „Breite“. Um einen funktionalen und adressatengerechten Text zu verfassen, muss der:die Schreibende sowohl über ein breites Wortwissen als auch über ein tiefes Verständnis verfügen, d. h. sie oder er muss ein geschriebenes Wort nicht nur wiedererkennen, sondern er oder sie muss in der Lage sein, die Bedeutung eines bestimmten Wortes zu definieren, um es produktiv in einem geschriebenen Text zu verwenden. Teil eines tiefen Wortschatzes sind Informationen über Form, Bedeutung und Verwendung eines Wortes in einem stark organisierten semantischen Netzwerk. Mit dem PPVT wird der tiefe Wortschatz nur unzureichend erfasst, da hier Informationen über Form und Verwendung nur marginal abgerufen werden müssen.

In einer weiterführenden Analyse wurde der Wortschatz in verfassten Texten, die der Beurteilung der Textproduktionskompetenz dienten (Beschreibungen eines Zoos), reanalysiert. Untersucht wurde das Vorkommen weniger gebräuchlicher Wörter. Hierzu wurde die Häufigkeit der geschriebenen Wörter mit der in „childLex“ beschriebenen Häufigkeit verglichen. Die Online-Datenbank basiert auf einem Korpus von Kinderbüchern (Schroeder et al., 2015). Angenommen wurde, dass Schüler:innen, die seltene Wörter verwenden, um ihre Ideen über einen Zoo zu beschreiben, über einen tiefen Wortschatz verfügen. Ein schneller Abruf auch von ungebräuchlichen Wörtern als Operationalisierung eines tiefen Wortschatzes sollte ein besserer Indikator für eine hohe Schreibflüssigkeit und Textproduktion darstellen als ein mit dem PPVT gemessener breiter Wortschatz.

Es zeigten sich geringe Interkorrelationen zwischen den Facetten des Wortschatzes, so dass davon ausgegangen werden kann, dass Tiefe und Breite unterschiedliche Dimensionen sind. Die Ergebnisse zeigten in die hypothetisierte Richtung: Sowohl für die Transkriptionsflüssigkeit als auch für die Formulierungsflüssigkeit zeigten sich in allen Klassenstufen höhere Zusammenhänge mit dem tiefen als mit dem breiten Wortschatz. Während die Korrelationen mit dem tiefen Wortschatz zwischen .24 und .57 lagen, korrelierte der breite Wortschatz nahezu nicht. Die Textqualität korrelierte höher mit dem breiten Wortschatz als mit dem tiefen (.40 bis .60 je nach Klassenstufe vs. .10 bis .33). Sowohl Tiefe als auch Breite tragen somit zum Schreiben bei. Im flüssigen Schreiben schwächeren

Schüler:innen fehlt möglicherweise ein tiefer Wortschatz. Eine Konsequenz hieraus wäre die Förderung eines in Bezug auf Form, Bedeutung und Verwendung tiefen Wortschatzes.

Im Vortrag geben wir einen detaillierten Überblick über den Zusammenhang von Wortschatz und Schreiben, zeigen das Strukturgleichungsmodell der ersten Studie, geben Einblicke in die theoretischen Annahmen für die Re-Analyse und präsentieren und diskutieren die Ergebnisse.

16. Historische Kompetenzen am Thema Transformationszeit nach 1990 standardisiert erfassen. Testentwicklung im Rahmen einer Interventionsstudie

Katharina Totter[1], Wolfgang Wagner[1], Christiane Bertram[2], Ulrich Trautwein[1]
[1]Universität Tübingen, Hector-Institut für Empirische Bildungsforschung, Germany; [2]Universität Konstanz, Binational School of Education, Fachbereich Geschichte und Soziologie

Theoretischer Hintergrund

Geschichtsunterricht soll das Handwerkszeug zum historischen Denken vermitteln, um historische Deutungen kritisch prüfen und ein reflektiertes Geschichtsbewusstsein ausbilden zu können (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016). Eine interdisziplinäre Arbeitsgruppe aus Geschichtsdidaktiker:innen und -wissenschaftler:innen sowie Bildungsforscher:innen der Universitäten Konstanz und Tübingen entwickelte in Zusammenarbeit mit dem Zentrum für Schulqualität und Lehrerbildung für eine groß angelegte randomisierte kontrollierte Feldstudie eine Unterrichtseinheit zum Thema Transformationszeit nach 1990 (Bertram, 2020). Diese sollte prozessbezogenen Kompetenzen historischen Denkens gemäß des FUER-Modells (Schreiber et al., 2006) gezielt fördern. Dementsprechend sollte der Zuwachs an im Rahmen der Unterrichtseinheit geförderten historischen Kompetenzen möglichst präzise und thematisch nah an der Intervention erfasst werden. Bisherige deutsche Inventare standardisierter Testaufgaben zum historischen Denken wie HiTCH (Trautwein et al., 2017) decken in der Unterrichtsstudie vermittelte Inhalte zur Transformationszeit nicht ab. An den bestehenden Aufgabenformaten im HiTCH-Inventar orientiert sollten Items mit neuem inhaltlichen Rahmen entsprechend den in der Studie geförderten Inhalten und Kompetenzen konzipiert werden.

Anhand des Posters soll die Entwicklung sowie Ergebnisse der Validierung dieses standardisierten Instruments zur Erfassung der prozessbezogenen historischen Kompetenzen am Thema der Transformationszeit dargelegt werden.

Fragestellung

Können die entwickelten Testinstrumente das historische Denken reliabel messen? Wie stehen sie mit bisherigen Schulleistungen, der Einschätzung der Relevanz von Geschichte, motivationalen Variablen zum Fach Geschichte und Wissensaufgaben zur Transformationszeit in Zusammenhang?

Methode

Die Operationalisierung hatte zum Ziel, standardisierte Testaufgaben zu entwickeln, die die in der Unterrichtseinheit geförderten Inhalte und historischen Kompetenzen möglichst genau abbilden. Schüler:innen lernten in der Intervention, die präsentierten Quellen zur Transformationszeit kritisch zu hinterfragen. Hierfür wurden Testaufgaben zur Methodenkompetenz konzipiert. Die Schüler:innen sollten außerdem dazu angeregt werden, über fremde und eigene Geschichtsbilder nachdenken und

ihre epistemologischen Überzeugungen zu schärfen, wofür Testaufgaben zur Reflexions- und Orientierungskompetenz entwickelt wurden. Im Hinblick auf den Einsatz des Tests in der Mittelstufe wurde insbesondere auf die Testökonomie und Zumutbarkeit geachtet. Der Test sollte einen hohen Aufforderungscharakter haben und zur Auseinandersetzung mit den Testfragen anregen, um möglichst viele Schüler:innen zur Bearbeitung zu motivieren.

Insgesamt wurden sieben Aufgabenformate (k_{Gesamt} = 58 Items) zu verschiedenen Bereichen der historischen Kompetenz überprüft. Im Fragebogen wurden zusätzlich ein Test zum historischen Wissen zur Transformationszeit, Skalen zur Relevanz von Geschichte und motivationale Inventare eingesetzt sowie Schulnoten und demographische Daten erhoben. Die insgesamt 55-minütige Testung wurde in 19 Klassen (davon n = 17 zehnte Klassen und n = 2 elfte Klassen) mit insgesamt 354 Schüler:innen in Baden-Württemberg durchgeführt.

Die Teilaufgaben zur historischen Kompetenz wurden mittels zweiparametrisch logistischer (2PL) Item Response Modellen untersucht. Als psychometrisches Auswahlkriterium für die Items wurde eine statistisch signifikant positive Diskrimination mit einem Kennwert von $a \geq 0.5$ festgelegt. Für die finale Itemauswahl (dabei spielten auch inhaltliche Kriterien eine Rolle), die unverändert in der Interventionsstudie eingesetzt wurde, wurde die Reliabilität ermittelt sowie Zusammenhänge mit Schulnoten untersucht. Weitere Zusammenhänge mit der Einschätzung der Relevanz von Geschichte, motivationalen Variablen zum Fach Geschichte und Wissensaufgaben zur Transformationszeit werden berichtet.

Zentrale Ergebnisse

Insgesamt konnten 39 der 58 Items in teilweise angepasster Form für die Testung der historischen Kompetenzen in der Interventionsstudie Verwendung finden. Anpassungen entfielen dabei hauptsächlich auf eines der Aufgabenformate (k = 24 Items). Für 25 Items aus den sechs verbleibenden Aufgabenformaten, die auf Basis der Analysen unverändert für die Interventionsstudie übernommen wurden, betrug die WLE-Reliabilität .69. Eine multiple Regression des WLE-Scores auf die Schulnoten in den Fächern Geschichte, Deutsch und Mathematik ergab lediglich für die Geschichtsnote (b = -.22, p = .024, n = 334) einen statistisch signifikanten Regressionskoeffizienten.

Diskussion/Ausblick

Die validierten Items können eine Erweiterung des bisherigen Pools an standardisierten Testaufgaben (vgl. Trautwein et al., 2017) zu historischen Kompetenzen darstellen. Ein Einsatz des Instruments in der Interventionsstudie soll weitere Erkenntnisse über die Güte der Aufgaben liefern.

17. Entwicklung und Erprobung eines Fragebogens zur Erfassung des Autonomiestrebens von Schüler:innen im Biologieunterricht

Nadine Großmann, Benjamin A. Tunnat, Tim Kirchhoff
Universität Bielefeld, Deutschland

Theoretischer und empirischer Hintergrund

Im Zuge der Inklusion sind Lehrpersonen zunehmend mit individuellen Unterschieden ihrer Schüler:innen konfrontiert, die sie diagnostizieren und in ihrer Unterrichtsgestaltung entsprechend berücksichtigen müssen. Diese individuellen Unterschiede müssen auch bei der Implementation

autonomieförderlicher Maßnahmen berücksichtigt werden. Autonomieförderliche Maßnahmen haben sich in diversen Studien als wirksame Mittel zur Förderung selbstbestimmter Motivationsqualitäten im Unterricht herausgestellt (bspw. Mittag et al., 2009; Großmann et al., 2022). Diese Maßnahmen wurden anhand des psychologischen Bedürfnisses nach Autonomie entwickelt, das in der Selbstbestimmungstheorie der Motivation (Ryan & Deci, 2017) verankert ist. Führen Individuen Handlungen freiwillig aus und nehmen in ihren Handlungen Wahlfreiheit sowie sich selbst als Verursacher dieser Handlungen wahr, ist ihr Bedürfnis nach Autonomie befriedigt (Reeve et al., 2003).

Erste Studien zeigen bereits auf, dass sich autonomieförderliche Maßnahmen je nach individuellen Voraussetzungen der Schüler:innen unterschiedlich auf ihre Motivationsqualität im Unterricht auswirken können (bspw. Großmann et al., 2022; Großmann & Wilde, 2020). Ein vielversprechender Ansatz zur Berücksichtigung individueller Unterschiede in der Untersuchung und Gestaltung von autonomieförderlichen Maßnahmen, der aktuell große internationale Aufmerksamkeit erhält, widmet sich dem Autonomiestreben (bspw. Van Assche et al., 2018; Wörtler et al., 2020). Mit dem Begriff Autonomiestreben werden individuelle Unterschiede im Wunsch, das eigene Bedürfnis nach Autonomie zu befriedigen, beschrieben (Van Assche et al., 2018). Da diese Variable vor allem im englischsprachigen Raum untersucht wird, fehlt im deutschsprachigen Raum ein hinreichend reliables und valides Messinstrument. Eine Entwicklung und Validierung eines deutschsprachigen fachspezifischen Messinstruments zur Erfassung des Autonomiestrebens von Schüler:innen wird daher im vorliegenden Projekt fokussiert.

Fragestellungen

Kann eine Ein-Faktor-Struktur des entwickelten Messinstruments bestätigt werden?

Kann das entwickelte Messinstrument das Autonomiestreben von Schüler:innen im Biologieunterricht reliabel messen?

Methode

Die Entwicklung des Messinstruments folgte dem TRAPD-Ansatz (Harkness, 2003). Hierfür wurden die Items des englischsprachigen Fragebogens von Van Assche et al. (2018) zum Autonomiestreben zunächst in einer Runde aus Fachdidaktiker:innen und Psycholog:innen übersetzt. Von einem bilingualen Muttersprachler mit Lehramtsausbildung wurden diese Items im Anschluss zurückübersetzt. In der o. g. Expertenrunde wurden alle Übersetzungen diskutiert und notwendige Anpassungen vorgenommen. Im Anschluss wurde eine Adjukatorin (Psychologin mit Erfahrung in der Fragebogenentwicklung und der Selbstbestimmungstheorie) einbezogen, mit der letzte offene Fragen und Diskussionspunkte geklärt wurden. Nach der Durchführung der Pilotstudie, die in diesem Abstract fokussiert wird, erfolgte die Auswertung der Daten und die Dokumentation des gesamten Prozesses.

135 Schüler:innen (62% weiblich; M_{Alter}=17.89 Jahre; SD_{Alter}=1.53 Jahre) der Sekundarstufe II nahmen an der Pilotstudie teil. Die Schüler:innen füllten nach ihrer schriftlichen Einwilligung einen Online-Fragebogen aus, der die entwickelten acht Items enthielt. Diese konnten die Schüler:innen anhand einer fünfstufigen Ratingskala für ihren Biologieunterricht bewerten. Zudem konnten sie

angeben, welche Items sie nicht bewerten konnten und warum sie diese Items nicht bewerten konnten.

Ergebnisse

Da das zugrundeliegende englischsprachige Messinstrument eine Ein-Faktor-Struktur aufweist (vgl. Van Assche et al., 2018), wurde auch für die entwickelten deutschsprachigen Items eine Ein-Faktor-Struktur angenommen. Die konfirmatorische Faktorenanalyse zeigt zufriedenstellende Kennwerte ($\chi^2(20)=27.40$, TLI=.96, CFI=.97, RMSEA=.06, SRMR=.06) für das aufgestellte Modell (vgl. Hu & Bentler, 1999; Kline, 2015; Moosbrugger & Schermelleh-Engel, 2012). Die Faktorladungen der Items liegen zwischen .47 und .72. Zudem zeigt der entwickelte Fragebogen eine gute interne Konsistenz (Cronbachs α=.82; DeVellis, 2012).

Fazit und Ausblick

In der Pilotstudie konnte die Ein-Faktor-Struktur des englischsprachigen Messinstruments für die entwickelten deutschsprachigen Items bestätigt werden. Anhand der Rückmeldungen der Schüler:innen wurden die Items sprachlich weiterentwickelt. Zudem wurden Validierungsinstrumente für die Hauptstudie ausgewählt, die zurzeit durchgeführt wird. Bspw. werden hier zur Prüfung der Kriteriumsvalidität die wahrgenommene Kompetenz und Autonomie im Biologieunterricht (Heissel et al., 2018; Wilde et al., 2009) sowie ihr Sachinteresse an Biologie (Frey et al., 2009) erhoben. In dieser Studie werden zur Kreuzvalidierung auch Lehramtsstudierende (Unterrichtsfach Biologie) miteinbezogen. Die Ergebnisse der Hauptstudie werden auf der Tagung vorgestellt.

18. Wann und wie dürfen Grundschulkinder im Unterricht mitbestimmen? Mitbestimmungsmöglichkeiten im Grundschulunterricht aus kinderrechtlicher und demokratiepädagogischer Perspektive

Leonora Gerbeshi, Sonja Ertl
Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Theoretischer und empirischer Hintergrund

Die Mitbestimmung von Kindern ist in der UN-Kinderrechtskonvention festgeschrieben (bmfsfj, 2019) und wird häufig mit Partizipation gleichgesetzt. Da Partizipation eher als Sammelbegriff zu verstehen ist (Reisenauer, 2020), wird Mitbestimmung im vorliegenden Vorhaben als Teil von Partizipation definiert und als kollektiver Prozess verstanden (Müller-Kuhn & Häbig, 2022), der von Selbstbestimmung abgegrenzt wird. Mitbestimmung wird im Folgenden in Anlehnung an Ertl, Martschinke & Grüning (2022) über drei Facetten operationalisiert: Informiert und gehört werden, Meinung und Abstimmung sowie Diskussion und Austausch. Dabei können sich die Facetten der Mitbestimmung sowohl auf inhaltliche Aspekte des Unterrichts, z. B. Unterrichtsthemen und Hausaufgaben, aber auch auf organisatorische Aspekte, wie Klassenregeln und Klassenzimmergestaltung, beziehen.

Aktuelle Befunde zur Mitbestimmung in Schule und Unterricht zeigen, dass im Grundschulbereich und im Rahmen des Unterrichts wenig Mitbestimmung gewährt wird (Müthing, Razakowski & Gottschling, 2018). Außerdem belegen Studien, dass (Grundschul-)Kinder und (Grundschul-)Lehrkräfte Mitbestimmungsmöglichkeiten unterschiedlich

wahrnehmen und Lehrkräfte die Mitbestimmung höher einschätzen als die Kinder selbst (Coelen & Wagener, 2011; Forde et al., 2018), wobei Mitbestimmung hier anders operationalisiert wird als in dem vorliegenden Beitrag. Darüber hinaus zeigen Studien, dass Mitbestimmungsmöglichkeiten in Zusammenhang zu unterschiedlichen Merkmalen, wie beispielsweise der Klassengröße (Bacher et al., 2007) oder dem Wissen der Kinder (Casas et al., 2018) stehen.

Forschungsdesiderat

Zum einen sind keine Studien bekannt, die Mitbestimmungsmöglichkeiten im Grundschulunterricht aus Kinder- und Lehrkräfteperspektive untersuchen und bezüglich verschiedener Ebenen und Facetten der Mitbestimmung differenzieren. Außerdem fehlen Untersuchungen, die darauf aufbauend Zusammenhänge zwischen Mitbestimmung und Merkmalen, wie beispielsweise dem Belastungsempfinden der Lehrkräfte, dem Mitbestimmungswunsch der Kinder oder der Klassengröße, erklären.

Forschungsfragen

(1) Gibt es einen Unterschied zwischen der Perspektive von Grundschulkindern und -lehrkräften bezüglich der Mitbestimmungsmöglichkeiten?

(2) Wie hängen Mitbestimmungsmöglichkeiten mit der Selbstbestimmung, dem Mitbestimmungswunsch, der institutionellen Mitbestimmung, dem Belastungsempfinden der Lehrkräfte und der Klassengröße zusammen?

Methode

Im Rahmen eines quantitativen Queerschnittsdesigns wurden Grundschullehrkräfte und Grundschulkinder der 3. Jahrgangsstufe mittels eines Fragebogens befragt (n^{Kinder}=551; $n^{Lehrkräfte}$=29).

Die wahrgenommenen Mitbestimmungsmöglichkeiten werden über selbst konstruierte und pilotierte Skalen entsprechend der Facetten der Mitbestimmung erfasst: Informiert und gehört werden, Meinung und Abstimmung sowie Diskussion und Austausch. Diese werden jeweils nach Unterrichts- und Klassenebene unterschieden. Die Reliabilitäten der sechs Skalen liegen alle im noch akzeptablen bis guten Bereich (α = .60 bis .86). Außerdem wurde etablierte Skalen eingesetzt, um die Selbstbestimmung (Röder & Kleine, 2007) sowie das Belastungsempfinden der Lehrkräfte (Wendt, Bos, Goy & Jusufi, 2017) zu erheben.

Die Perspektiven der Lehrkräfte und Kinder sollen mittels deskriptiver Daten und eines U-Tests miteinander verglichen werden. Der Zusammenhang von Mitbestimmungsmöglichkeiten und Merkmalen wie dem Belastungsempfinden der Lehrkräfte oder der Klassengröße soll mittels Korrelationsanalysen untersucht werden.

Ergebnisse

Erste Ergebnisse zeigen, dass Grundschullehrkräfte Mitbestimmungsmöglichkeiten im Unterricht, abgesehen von der Skala *Diskussion auf Unterrichtsebene*, auf allen Ebenen (Unterrichts- und Klassenebene) und Facetten signifikant höher einschätzen als die Grundschulkinder. Die Ergebnisse bezüglich der zweiten Forschungsfrage werden momentan ausgewertet und auf der Tagung präsentiert.

Die Ergebnisse bieten einen entscheidenden Beitrag zur Partizipationsforschung und dienen der differenzierten Erklärung von Mitbestimmungsmöglichkeiten von Grundschulkindern im Unterricht.

19. Gesundheitsförderung in Schulen – eine Vergleichsstudie zwischen Malaysia und der Schweiz

Simona Betschart
Schwyz University of Teacher Education, Switzerland

Die Schule zählt spätestens seit der Ottawa-Charta (WHO, 1986) als zentraler Ort für Gesundheitsförderung. Dabei gilt insbesondere ein ganzheitlicher Ansatz, bei dem sämtliche Mitglieder einer Schule und die Schule als Ganzes miteinbezogen werden, als erstrebenswert (WHO Expert Committee, 1997). Schulleiter*innen werden als zentrale Schlüsselfiguren für die Implementation von gesundheitsfördernden Massnamen betrachtet (Dadaczynski et al., 2021; Gieske & Harazd, 2009; Zumbrunn et al., 2016), wobei wenig darüber bekannt ist, wie genau sie Gesundheitsförderung steuern und welche Eigenschaften für die damit verbundenen Führungspraktiken wichtig sind (Dadaczynski et al., 2022). Auf Ebene der Schulleitung können verschiedene kontextuelle und personelle Faktoren, wie bspw. das Geschlecht, die Einstellung gegenüber Gesundheitsförderung oder Health Literacy von Schulleiter*innen von Relevanz für die Implementation von Gesundheitsförderung sein (Clarke et al., 2015; Dadaczynski et al., 2020; Dadaczynski & Paulus, 2015).

Dieses vergleichende Projekt beschäftigt sich damit, inwiefern Schulleiter*innen aus Malaysia und der Schweiz während der Covid-19-Pandemie schulische Gesundheitsförderung implementieren konnten. Der Vergleich zwischen der Schweiz und Malaysia ist von Interesse, da sich die beiden Länder in Bezug auf die Educational Governance unterscheiden. In der Schweiz können Schulleiter*innen lokale Prioritäten setzen, müssen sich aber gleichzeitig an kantonale Richtlinien halten. Im Gegensatz dazu, hat Malaysia ein deutlich zentralisiertes System: Im Bildungsbereich sind formale Ziele der Schulen klar vorgegeben. Ungeachtet dieser Unterschiede standen beide Länder im Zusammenhang mit der Covid-19-Pandemie vor ähnlichen Herausforderungen. Dennoch beeinflusst der nationale und lokale Kontext, inwiefern Schulleiter*innen ermächtigt und befähigt werden (vgl. Sears & Marshall, 1990), auf Herausforderungen wie die Covid-19-Krise zu reagieren und entsprechend dem lokalen Kontext Anpassungen vorzunehmen.

Im Juni 2021 fand eine Online-Befragung im statt, woran N = 1058 Schulleiter*innen aus beiden Ländern teilnahmen. Die Umsetzung der Gesundheitsförderung wird anhand der Survey of School Promotion of Emotional and Social Health (SSPESH)-Skala (Dix et al., 2019) gemessen, die auf dem Rahmenkonzept für gesundheitsfördernde Schulen (HPS) der Weltgesundheitsorganisation (WHO) basiert (Langford et al., 2015) und das Ausmass misst, in dem eine Schule Massnahmen und Praktiken zur Förderung der emotionalen und sozialen Gesundheit der Schüler*innen umgesetzt hat. Erste Ergebnisse aus Regressionsanalysen und Simple Slope Analysen zeigen, dass Einstellungen gegenüber Gesundheitsförderung, die Zusammensetzung der Schülerschaft sowie die Schulstufe in beiden Ländern relevante Faktoren für die Implementation von Gesundheitsförderung sind. Die Einstellungen gegenüber Gesundheitsförderung zeigen zudem einen unterschiedlichen Effekt in den

beiden Ländern. Es scheint, dass eine positive Einstellung gegenüber Gesundheitsförderung in der Schweiz von höherer Relevanz ist als in Malaysia.

Im Anschluss and die quantitativen Erhebungen fand im Mai 2022 eine qualitative Follow-Up Studie statt, bei der Interviews mit Schulleiter*innen in der Schweiz und Malaysia geführt wurden. Erste Ergebnisse liefern weitere Hinweise zu möglichen Faktoren auf nationaler und lokaler Ebene, welche die Praktiken der Gesundheitsförderung beeinflussen.

20. Die Bedeutung von Wissen für eine Entscheidung zum eigenen Gesundheitsverhalten

Moritz Klatthaar, Matthias Wilde, Melanie Basten
Universität Bielefeld, Germany

Theoretischer Hintergrund

Nicht zuletzt die Corona-Pandemie hat gezeigt, dass Entscheidungen bezogen auf das eigene Gesundheitsverhalten, die Auswirkungen auf die gesamte Gesellschaft haben, von einer ausreichenden naturwissenschaftlichen Grundbildung (Bennett et al., 2007) abhängen (Basten & Großmann, 2022). Dasselbe gilt für das Thema Organspende, bei dem die individuelle Entscheidung für oder gegen eine Spende gesamtgesellschaftlich hilft, die Diskrepanz zwischen benötigten und verfügbaren Spenderorganen zu verringern (Zimmering & Caillet-Brillet, 2021). Die Einstellung zur Organspende ist in Deutschland sehr positiv, geht aber häufig nicht mit der Dokumentation des eigenen Willens einher (ebd.). Die wichtige Rolle von Wissen für die Einstellungsbildung und Entscheidungsfindung hinsichtlich Organspende ist bereits mehrfach festgestellt worden. Allerdings unterscheiden sich die bisherigen Befunde in den berichteten Zusammenhängen (vgl. Horton & Horton, 1991; Morgan et al., 2002; Reubsaet et al., 2001). Gleichwohl ist das Thema Organspende bereits ab der Mittelstufe für die schulische Gesundheitsbildung relevant, da die Dokumentation des eigenen Willens auf einem Organspendeausweis ab dem Alter von 16 Jahren beworben wird (TPG, 2007).

Fragestellung

Welchen Einfluss hat das Wissen von Schüler:innen auf Einstellung, Verhaltensintention und Verhalten bei einem Gesundheitsthema mit gesamtgesellschaftlichen Auswirkungen (Organspende)?

Methode

Stichprobe

An der Studie nahmen 207 Schüler:innen teil. Der Großteil war zwischen 16 und 20 Jahren alt, nur sehr wenige waren älter (M = 17.13, SD = 1.30). Die analysierten Daten beziehen sich auf eine Vortestung vor einer Unterrichtseinheit zum Thema Organspende.

Instrumente

Die Schüler:innen füllten einen Fragebogen mit bipolaren Items mit siebenstufiger Antwortskala (0-6) zu ihrer Einstellung (neun Item; α = .92; vgl. Basten, 2013) und Verhaltensintention (sechs Items; α = .75; vgl. ebd.) aus. Das Wissen wurde mittels 16 offener Fragen erhoben (Interrater-Übereinstimmung der Auswertung von mindestens 90%), zwei davon waren entscheidungsbezogen. Das Verhalten

wurde als Besitz eines Organspendeausweises definiert und als Selbstauskunft mit den Antwortmöglichkeiten Ja und Nein erfragt.

Datenanalyse

Der angenommene Zusammenhang zwischen Wissen und Verhalten wurde mithilfe des Makros PROCESS 3.5 (Hayes, 2018) für IBM SPSS Statistics 26 getestet. Die abhängige Variable in unserem Modell war das Verhalten (Y). Die unabhängigen Variablen waren entscheidungsbezogenes Wissen und allgemeines Wissen (X). Als Mediatoren für die Regression von X auf Y wurden Einstellung (M1) und Intention (M2) gewählt. Schulklassen wurden dummy-kodiert und als Kovariaten eingebunden.

Ergebnisse und ihre Bedeutung

In einer ersten multiplen linearen Regression waren entscheidungsbezogenes Wissen (p = .171) und allgemeines Wissen (p = .208) die unabhängigen Variablen und Einstellung die abhängige Variable. Es konnte kein signifikanter Effekt der beiden Wissensvariablen auf Einstellung gefunden werden. In der zweiten multiplen linearen Regression waren entscheidungsbezogenes Wissen (p = .817), allgemeines Wissen (B = 0.54, $SE\ B$ = 0.06, p < .001) und Einstellung (B = 0.96, $SE\ B$ = 0.34, p = .005) die unabhängigen Variablen und Intention die abhängige Variable. Eine logistische Regression setzte sich aus entscheidungsbezogenem Wissen (B = 2.09, $SE\ B$ = 0.55, p < .001), allgemeinem Wissen (p = .553), Einstellung (p = .211) und Intention (B = 0.87, $SE\ B$ = 0.34, p = .011) als unabhängigen Variablen und dem Verhalten als abhängiger Variable zusammen. Die indirekten Effekte von entscheidungsbezogenem Wissen (95% CI = -0.05, 0.49) und allgemeinem Wissen (95% CI = -0.18, 1.14) auf Verhalten waren nicht signifikant.

Insgesamt zeigt sich also, dass Wissen zum Thema Organspende in einer nicht zuvor unterrichteten Stichprobe keinen Einfluss auf die Einstellung hat. Das allgemeine Wissen hat einen direkten Effekt auf die Verhaltensintention, während das spezifische entscheidungsbezogene Wissen einen direkten Einfluss auf das Verhalten (Besitz eines Organspendeausweises) hat. Auf dem Poster werden auch die Ergebnisse nach der Unterrichtung der Schüler:innen dargestellt und vergleichend vor dem Hintergrund der Theory of Planned Behavior (Fishbein & Ajzen, 2010) und der Gesundheitsbildung (bspw. Schaal et al., 2020) diskutiert.

21. Ein internationaler Vergleich von Lernergebnissen und sozialen Ungleichheiten in der schulischen politischen Bildung

Raphaela Schlicht-Schmälzle
Leibniz Institut für Bildungsforschung und Bildungsinformation, Deutschland

Fragestellung

Welche Lernziele erreicht schulische politische Bildung im internationalen Vergleich und welche sozialen Ungleichheiten entstehen dabei? Diese Studie untersucht, welches politische Grundlagenwissen Schülerinnen und Schüler in OECD Ländern erwerben und wie stark der Erwerb des politischen Grundlagenwissens von der sozialen Herkunft abhängt. Darauf aufbauend werden bildungssystemischen Faktoren identifiziert, die Erklärungen für die internationalen Unterschiede

liefern können. Für die Analysen werden Daten der International Civic and Citizenship Education Study (ICCS) 2016 (Köhler et al. 2018) genutzt.

Theoretischer Hintergrund

Zahlreiche Schulleistungsstudien haben Lernergebnisse international verglichen und somit international einen Wettbewerb der Outcome-Steuerung in der Bildungspolitik eingeleitet (Ender 2021). Ein häufiger Vorwurf gegenüber dieser Form der bildungspolitischen Steuerung ist ein verengter Blick auf das, was Schule leisten soll (Casey 2004; Fuchs 2003; Rutkowski 2007). Neben dem Erwerb von Wissen und Kompetenzen in einzelnen Fächern, hat Schule in den meisten Ländern auch einen Auftrag im Hinblick auf gesellschaftspolitische Lernziele (z.B. Galston 2001). Dennoch wissen wir bisher wenig darüber, welche kognitiven, behavioralen und ethischen Lernziele in Form von politischer Bildung in der Schule vermittelt werden und wie unterschiedliche Bildungssysteme diese Lernergebnisse erzielen. Die International Civic and Citizenship Education Study (ICCS) liefert eine solche erweiterte Perspektive, in dem sie sich einer Vielzahl an Lernergebnisse politischer Bildung von 13-jährigen im internationalen Vergleich widmet. Die vorliegende Studie konzentriert sich auf das erworbene gesellschaftspolitische Grundlagenwissen. Unterschiede im internationalen Vergleich legen nahe, dass diese durch makrogesellschaftliche Faktoren der Bildungssysteme zustande kommen (z.B. Schmidt et al. 2015). Basierend auf bestehender Literatur (z.B. Conover & Searing 2000; Galston 2001; Gamoran & Mare 1989; Niemi & Junn 1998; Wang et al. 1990), werden in dieser Studie eine Reihe von bildungssystemischen Faktoren daraufhin untersucht, ob sie Erklärungen für die internationalen Unterschiede in den Lernergebnissen und den sozialen Ungleichheiten liefern können.

Methode

Für die Analysen werden Daten der ICCS 2016 für die darin enthaltenden OECD-Länder analysiert. Anhand von Mehrebenenanalysen (vgl. Schlicht-Schmälzle et al. 2010) wird untersucht inwiefern sich der politische Wissenserwerb zwischen den OECD-Ländern unterscheidet, ob soziale Herkunftsfaktoren den politischen Wissenserwerb beeinflussen und ob sich diese Einflüsse ebenfalls zwischen den Länder unterscheiden. Zur Messung des politischen Grundlagenwissens dienen die für alle Teilnehmenden enthaltenen „Plausible Values“. Soziale Herkunft wird durch den elterlichen Bildungsstand (ISCED) gemessen. Darüber hinaus werden auch Indikatoren des Migrationshintergrunds und des Geschlechts in die Modelle integriert. In einem weiteren Schritt werden makrosystemische Faktoren in die Modelle integriert um zu analysieren, wie diese mit dem politischen Grundlagenwissen von Jugendlichen unterschiedlicher sozialer Herkunft zusammenhängen.

Ergebnisse

Die Analysen zeigen im internationalen Vergleich deutliche Unterschiede im Erwerb politischen Grundlagenwissen. Der Einbezug individueller sozialer Faktoren zeigt, dass die Bildungsherkunft den politischen Wissenserwerb positiv beeinflusst, dass Mädchen mehr politisches Grundlagenwissens haben als Jungs und dass Jugendliche mit Migrationshintergrund weniger politisches Grundlagenwissen erworben haben als ihre Vergleichsgruppe ohne Migrationshintergrund. Jedoch

unterscheidet sich das Ausmaß des Einflusses sozialer Faktoren auch zwischen Ländern. Diese Unterschiede weisen darauf hin, dass makrogesellschaftliche Faktoren im Bildungssystem den Erwerb von politischer Bildung und deren ungleiche Verteilung beeinflussen.

Tatsächlich zeigen sich Zusammenhänge zwischen bildungssystemischen Größen und dem Erwerb des individuellen politischen Grundlagenwissens. Die Größe des Privatschulsektors, die Gliederung in der Sekundarstufe, die Integration von politischer Bildung in weitere curriculare Bereiche, eine Kultur der Diskussion in der Schule und die Qualifikation der Lehrkräfte weisen Zusammenhänge mit dem Erwerb politischen Grundlagenwissens und teilweise auch mit den sozialen Ungleichheiten auf. Die Ergebnisse liefern erste Hinweise darauf, wie bildungssystemische Rahmenbedingungen die politische Bildung in allen gesellschaftlichen Gruppen stärken kann. Um die kausalen Mechanismen dieser Zusammenhänge nachzuvollziehen, ist weitere Forschung auf Ebene der Klassenzimmer und Schulen notwendig.

22. Quantitative Erprobung eines kompetenzorientierten Leistungstests auf unternehmerisches Wissen und Denken. Diagnostik als Basis effektiver Entrepreneurship Education (EE) in der Schule.

Steffen Spitzner[1], **Thomas Retzmann**[2]
[1]Universität Duisburg-Essen, Deutschland; [2]Universität Duisburg-Essen, Deutschland

Entrepreneurship als theoretischer und gesellschaftlicher Hintergrund

Entrepreneurship wird neben dessen ökonomischer Bedeutung das Potenzial zugeschrieben, zur Entfaltung der Gesellschaft beizutragen (Fritsch 2019). Die aktuellen Herausforderungen infolge der demografischen Entwicklung, des Klimawandels und der politisch-ökonomischen Krisen hinsichtlich Versorgungs- und Energiesicherheit erfordern neuartige Problemlösungen vonseiten innovativer Akteure jedweder Couleur (Faltin 2018). Viele Entrepreneure treten an, um innovative und bisweilen disruptive Lösungen für die ökonomischen, sozialen und ökologischen Probleme der Zeit zu entwickeln und gründen Startups. Da diese sich typischerweise in besonders dynamischen Umfeldern bewegen, müssen sie lernen, mit Unsicherheit kompetent umzugehen und Risiken angemessen abzuschätzen. Individuelle Ressourcen wie Selbstständigkeit, Leistungsmotivation, Risikoneigung, Innovationsfähigkeit und Verantwortungsübernahme erscheinen relevanter denn je (Retzmann & Hausmann 2012).

Im vorliegenden Forschungskontext wird, in Übereinstimmung mit der Entrepreneuship-Forschung, die Innovation als das Bestimmungsmerkmal von Entrepreneurship angesehen und die Risikoübernahme als Abgrenzungsmerkmal zum Intrapreneurship. Beide Formen der selbstständigen Erwerbstätigkeit verbindet der innovativ-unternehmerische Prozess, der den Gegenstand maßgeblich charakterisiert. Weil viele der typischen Fähigkeiten und Eigenschaften von Entrepreneuren auf Intrapreneure übertragbar sind, ist der Nutzen einer EE nicht auf die unternehmerische Selbstständigkeit oder die Unternehmensgründung limitiert (Retzmann & Seeber 2019). Die subjektorientierte EE kann zu einer generellen Beschäftigungsfähigkeit beitragen, weil diese Fähigkeiten und Eigenschaften zunehmend von Arbeitnehmer*innen erwartet werden. Lernende sollen daher befähigt werden, erwerbstätig geprägte Situationen der unternehmerischen und

beruflichen Selbstständigkeit (zukünftig besser) zu bewältigen sowie die Gesellschaft mündig, verantwortungsvoll und aktiv mitzugestalten (Bacigalupo et al. 2016; Kirchner & Loerwald 2014).

Jedoch ist das Gründungsgeschehen in Deutschland im internationalen Vergleich gering ausgeprägt und langfristig rückläufig; die EE wird hierzulande seit Jahren als schlechteste entrepreneuriale Rahmenbedingung identifiziert (Metzger 2022; Sternberg et al. 2022). Um schulische und außerschulische Maßnahmen der EE zielführend und effektiv zu gestalten, bedarf es empirisch belastbarer Rahmenmodelle und darauf fußender, standardisierter Messinstrumente, die in der Lage sind, den Learning Outcome zu erheben, der zu dem o. g. gesamtwirtschaftlichen und gesellschaftlichen Impact führen soll. In Ermangelung solcher Instrumente liegen derzeit kaum wissenschaftliche Erkenntnisse zur Effektivität der scheinbar zunehmenden Maßnahmen der schulischen und außerschulischen EE (schulische Differenzierungskurse, regionale und nationale Wettbewerbe etc.) vor. Für die non-kognitive Kompetenzfacette finden sich erste Ansätze (Grewe 2020; Kriedel 2017); für die kognitive Dimension des Wissens und Denkens kann jedoch – bisher – nicht auf entsprechende Instrumentarien zurückgegriffen werden. Allerdings spielen fachbezogene Fähigkeiten und fachbezogenes Wissen eine zentrale Rolle, wenn die Entwicklung fachübergreifender Kompetenzen das Vorhandensein gut ausgeprägter Fachkompetenzen voraussetzt (Klieme et al. 2007; Weinert 2002). Es erfordert daher Messinstrumente, die eine fundierte Diagnose und gezielte Förderung ermöglichen.

Fragestellung, Vorgehen, Ergebnisse und deren Bedeutung

Auf Basis des eigens konzipierten Rahmenmodells und qualitativer Überprüfungsschritte mit unterschiedlichen Expertengruppen wurde der Fragestellung nachgegangen, ob sich unternehmerisches Wissen und Denken als eigenständiger Gegenstandsbereich konstruieren und inhaltsvalide als eindimensionales Konstrukt objektiv, differenziert, trennscharf, zuverlässig und fair messen lässt. Um sie zu beantworten wurden die klassische und probabilistische Testtheorie kombiniert und die 1-PL-Rasch-Skalierung inkludiert.

Im Mittelpunkt des Vortrags stehen die Ergebnisse der quantitativen Erprobung eines neu entwickelten Leistungstests auf unternehmerisches Wissen und Denken. Diese erfolgte 2021 bis 2022 mit über 700 Schüler*innen berufsbildender Schulen in Nordrhein-Westfalen. Die Itemselektion samt definierter Ein- und Ausschlussgrenzen rekurriert sich zunächst auf die Parameter (Person-Wright-Map), korrigierte Trennschärfe sowie die Zuverlässigkeit des Tests (EAP/PV-Reliabilität). Die Eindimensionalität wurde anhand relevanter Itemfit-Werte (Infit & Outfit) überprüft, ehe unter Einbezug verschiedener Teilungskriterien Analysen zum Differential Item Functioning erfolgten, um die Testfairness zu gewährleisten. Die für eine aussagekräftige pädagogische Diagnose entwickelten Kompetenzstufen auf Basis der Item-Response-Theorie dienen der kriteriumsorientierten Interpretation der individuell erzielten Testwerte.

Abschließend erfolgt neben einer kritischen Würdigung ein kurzer Ausblick auf den nächsten Schritt der Testnormierung, um durch weitergehende empirische Bildungsforschung (vor allem Interventionsstudien) zunehmende Evidenz für die effektive Gestaltung von Maßnahmen der EE in der Bildungspraxis zu generieren.

23. Zur richtigen Zeit am richtigen Ort? Zur Rolle der Sitzpartnerschaft im Peereinfluss bei der Entwicklung sozialer Kompetenzen

Corinna Hank, Philipp Nicolay, Christian Huber
Bergische Universität Wuppertal, Germany

Im schulischen Setting gewinnt der Einfluss durch Peers mit zunehmendem Alter an Bedeutung. Insbesondere für antisoziales Verhalten gilt der Einfluss der Klassengemeinschaft als bestätigt (z.B. Müller, Hofmann & Arm, 2017). Auch für prosoziales Verhalten zeigen sich Befunde, die den positiven Einfluss einer prosozialen Peergemeinschaft auf die individuelle Entwicklung prosozialen Verhaltens unterstützen (z.B. Busching & Krahé, 2020). Für den schulischen Alltag stellt der Peereinfluss an einigen Stellen eine Herausforderung für Lehrkräfte dar, könnte aber auch zur Unterstützung der Entwicklung erwünschter Kompetenzen wie beispielsweise sozialer Kompetenzen genutzt werden. Vor dem Hintergrund, dass neben einer inhaltlichen Vermittlung auch soziales Lernen in die Verantwortung von Klassenlehrkräften gelegt wird (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2012), erscheint es notwendig, Bedingungen für eine zeitgleiche Förderung von Inhalt und sozialen Kompetenzen zu optimieren. Der Peereinfluss stellt hier einen Mechanismus dar, der eine zeitlich ökonomische sowie implizite Unterstützung sozialer Lernprozesse ermöglichen könnte. Eine Option den Peereinfluss zu steuern, ist durch die Steuerung der Nähe zu bestimmten Peers gegeben. Während eine Vielzahl an Studien den Ausprägungsgrad der relevanten Kompetenz der Klassengemeinschaft als Umgebungsfaktor der Individuen berücksichtigt, fehlen Untersuchungen zum Einfluss der unmittelbaren Sitzpartnerschaft im Unterricht. Dieser Beitrag hat es sich zum Ziel gemacht, diese Lücke zu schließen und den Einfluss der unmittelbaren örtlichen Peerumgebung auf die Entwicklung sozialer Kompetenzen in den Fokus zu nehmen. Darüber hinaus wird hier explorativ untersucht, inwiefern sich eine Veränderung der Sitzordnung in einer Klasse auf ebendiese Veränderung auswirkt. Vor diesem Hintergrund untersucht der Beitrag folgende Fragestellungen:

1.) Sind Schüler:innen ihren Sitznachbar:innen in ihren sozialen Kompetenzen ähnlicher als anderen Kindern der Klasse?

2.) Verändern sich selbst wahrgenommene und durch die Lehrkraft wahrgenommene soziale Kompetenzen im Sinne einer Angleichung an einen neuen Sitznachbarn, wenn sich die Sitzordnung verändert wird.?

Mithilfe von *N* = 202 Schülerinnen und Schülern aus 14 Klassen und damit 101 Dyaden wurden diese Fragestellungen untersucht. Die 14 Klassen wurden in eine Interventions- und eine Kontrollgruppe unterteilt und durch zwei Erhebungszeitpunkte im Abstand von fünf Wochen begleitet. Nach einer ersten Erhebung wurde in sieben Interventionsklassen die Sitzordnung zufällig neu zusammengesetzt. Hier galt die Einschränkung, dass alte Sitznachbarschaften nicht erneut bestehen durften. Die sieben Klassen der Kontrollgruppe behielten die ursprüngliche Sitzordnung bei. Zu den Erhebungszeitpunkten wurden neben der Sitzordnung die sozialen Kompetenzen der Schülerinnen Schüler sowohl aus Perspektive der Schülerinnen und Schüler als auch der Lehrkräfte erfasst. Zu diesem Zweck wurde die deutsche Übersetzung der Social Skill Improvement System Rating Scales SSIS RS (Gresham & Elliott, 2008; Hank & Huber, submitted) eingesetzt. Die Daten werden

mehrebenanalytisch mit dem R-Paket lme4 (Bates, Mächler, Bolker & Walker, 2015) ausgewertet. Es werden die Level der Messzeitpunkte, der Dyaden sowie der Individuen berücksichtigt.

Dieser Beitrag präsentiert Ergebnisse zum Einfluss der Sitzordnung auf die Entwicklung sozialer Kompetenzen. Damit diskutiert er auch die Möglichkeit der systematischen Integration einer gesteuerten Zuordnung der Sitzordnung sowie Unterstützungsoptionen, die die Sitzordnung vor dem Hintergrund verschiedener Förderziele bieten könnte. Die Ergebnisse werden vor dem Hintergrund des kurzen Beobachtungszeitraums sowie seiner lediglich ergänzenden Rolle in der Förderplanung reflektiert. Ein Bezug zu anschlussfähigen Unterrichtsmethoden, die Kontakte ebenfalls in den Fokus nehmen (wie z.B. das kooperative Lernen; (Hank, Weber & Huber, 2022; Jurkowski & Hänze, 2012) wird hergestellt.

24. Effects of a Spatial Thinking Intervention for Talented Elementary Students: A Randomized Controlled Field Trial.

Fabienne Kremer[1], Franziska Rebholz[2], Lucas Stark[1], Lisa Bardach[1], Jessika Golle[1], Benjamin Nagengast[1], Ulrich Trautwein[1]

[1]University of Tübingen, Germany; [2]Reutlingen University, Germany

Early spatial thinking skills predict STEM career choices and increase the likelihood of earning advanced STEM degrees (Wai, Lubinski, & Benbow, 2009). Nonetheless, the promotion of spatial thinking is neglected in school curricula (National Research Council, 2006). This might especially pose an obstacle for girls' success in STEM fields, as their disadvantages on spatial tasks have been observed in several yet not all studies (e.g., Linn & Petersen, 1985; Maeda & Yoon, 2013; Lauer, Yhang, & Lourenco, 2019). Spatial thinking skills are malleable through spatial interventions (Uttal et al., 2013). Early spatial interventions are called for to prevent students from getting overwhelmed by STEM content knowledge in later years of education (e.g., Newcombe, 2010; Lauer, Yhang, & Lourence, 2019).

We developed a spatial thinking intervention for first and second graders. It is part of the Hector Children's Academy Program (HCAP), an extracurricular enrichment program for talented elementary school children. The intervention addressed children's perspective taking and mental rotation skills in eight course units, each lasting 90 minutes. Overall, the intervention included recurring elements of embodied actions, imagination tasks, cooperative learning situations, and reflection about observations, whereas the task complexity increased systematically throughout the intervention. Following this structure, we aimed to develop the use of spatial language, move children from embodied to imaginary spatial thinking, create awareness of regularities and strategies, and gradually improve their ability to take another perspective and rotate objects before their mind's eye.

In the current study, a multisite cluster randomized controlled field trial with pre- and post-test measures was conducted at 7 HCAP sites (N = 97) to evaluate the intervention's effects on children's spatial thinking skills as well as on their motivation for spatial thinking. The intervention effects were assessed with an adapted version of the Perspective-Taking Test for Children (Frick, Möhring, & Newcombe, 2014) and subtests of 2D and 3D mental rotation tasks (Neuburger, Jansen, Heil, & Quaiser-Pohl, 2011). We preregistered the data analysis strategy on the Registry of Efficacy and Effectiveness Studies (REES, Registry ID: 10260.1v1) and performed multilevel regression analyses.

The analyses revealed significant positive effects of the treatment for two out of three spatial tasks: perspective taking (β = 0.312, *p* = .010) and one subtest of mental rotation (β = 0.383, p = .002). Participants' motivational indicators for spatial thinking were not significantly affected by the intervention program (self-concept of ability for spatial thinking: β = 0.240, *p* = .202, intrinsic value: β = -0.259, *p* = .117, attainment value: β = -0.302, *p* = .112). The intervention program's potential differential effects were additionally explored, showing, for instance, no significant differential effects for boys and girls on any of the spatial tasks (all p-values > .05). These results indicate that, after attending the course, children in the intervention group were better able to solve perspective taking and mental rotation tasks than those in the control group, while motivation for spatial thinking did not differ between the two groups, and that the intervention was equally beneficial for both sexes. The intervention thus represents a promising way to promote spatial thinking skills already early in elementary school.

25. Förderung situativer, aufgabenbezogener Emotionen von Studierenden durch eine Value-Intervention

Monika-Sybille Wolff
Universität Passau, Deutschland

Emotionen spielen eine wesentliche Rolle für Motivation, Leistung und erfolgreiches Lernen (vgl. z.B. Pekrun, Goetz, Titz & Perry, 2002; Pekrun, 2006; Pekrun & Perry, 2014). Gleichzeitig lassen sich aber über die Schulzeit hinweg eher ungünstige Entwicklungsverläufe erkennen: Während positive Leistungsemotionen wie Lernfreude abnehmen, werden negative Emotionen wie Prüfungsangst oder Langeweile häufiger erlebt (z.B. Vierhaus, Lohaus & Wild, 2016). Eine Analyse des Forschungsstandes zu Emotionsinterventionsmaßnahmen lässt eine Vielzahl an theoretisch abgeleiteten Möglichkeiten zur Förderung des emotionalen Erlebens in Lern- und Leistungssituationen finden, um diesem Abwärtstrend und den damit verbundenen negativen Folgen entgegenzuwirken. Empirisch überprüfte Interventionsmaßnahmen liegen aber bislang nur in überschaubarer Anzahl und primär für Prüfungsangst vor (z.B. Hall et al., 2007; Hembree, 1988). Trotz erster Erfolge, stellen sich viele dieser Interventionen in der Praxis als schwierig umsetzbar heraus (Brandenberger, Hagenauer & Hascher, 2018) oder verlieren nach Beendigung der Intervention ihre Effektivität (z.B. Hagenauer, Klaß & Gläser-Zikuda, 2016). Um eine positive emotionale Entwicklung zu begleiten, bedarf es demnach Interventionsmaßnahmen, die langfristig und unkompliziert im Unterricht integriert werden können.

Ausgehend von diesen Überlegungen wird in der vorliegenden Studie ein vielversprechender Ansatz aus der *utility-value*-Interventionsforschung auf die Förderung von Leistungsemotionen übertragen und die Umsetzung an einer Interventionsmaßnahme erprobt. Da *value*-Interventionen (z.B. Harackiewicz, Tibbetts, Canning & Hyde, 2014) auf eine Verbesserung der subjektiven Kontroll- und Wert-Einschätzungen von Leistungsaktivitäten durch direkte bzw. indirekte Prompts abzielen, sollten diese nicht nur Auswirkungen auf Motivation, sondern - gemäß Kontroll-Wert-Theorie (Pekrun, 2000, 2006) - über die Beeinflussung der Kontroll- und Wert-*Appraisals* auch auf Leistungsemotionen haben. Folglich wird in der vorliegenden Studie folgender Forschungsfrage nachgegangen: ***Welche***

Effekte haben wertinduktive Verarbeitungsprompts beim Bearbeiten von Lernaufgaben auf das durch Valenzkognitionen mediierte situative emotionale Erleben von Studierenden?

Aufgrund der fluktuierenden Natur von Emotionen (Ahmed, van der Werf, Minnaert & Kuyper, 2010) und der intra-psychologischen Prozesse, die der Emotionsentstehung inhärent sind (Pekrun, 2006), wird ein situativer und intraindividueller Ansatz in der Studie verfolgt. In einem experimentellen *within-subjects*-Design erhielten Lehramtsstudierende (*N* = 312) aus zwei Online-Seminaren (zum Thema „Lernemotionen“ und „Klassenführung“) über jeweils ein Semester verteilt drei *utility-value*-Treatments in Form einer Reflexionsaufgabe, die der eigentlichen Lernaufgabe vorgeschaltet war. Über direkte Prompts in Form von Zitaten junger Erwachsener, die über persönliche Erfahrungen berichteten, wurde den Studierenden zunächst die Bedeutsamkeit der in der Aufgabe angestrebten Kompetenzen vermittelt. Anschließend wurden die Studierenden dazu angeleitet, die Zitate hinsichtlich ihrer persönlichen Relevanz eigenständig zu evaluieren (vgl. Gaspard, Häfner, Parrisius, Trautwein & Nagengast, 2017). Das situative emotionale Erleben während des Bearbeitens der Lernaufgabe sowie die Valenz- und Kontrollkognitionen wurden mittels Fragebogen nach Abschluss der Lernaufgabe an Messzeitpunkten, an denen das Treatment stattfand, sowie zu drei weiteren Messzeitpunkten (alternierend) ohne Treatment erfasst.

Zur Untersuchung intraindividueller Unterschiede und Veränderungen im situativen Emotionserleben wurden Mehrebenenanalysen mit ‚*Treatment*‘ als zeitvariiernede Kovariate durchgeführt, um Diskontinuität in den Messwerten im Verlauf der Zeit zu modellieren (Singer & Willett, 2003). Im *Lernemotionen*-Seminar konnte ein positiver Treatmenteffekt auf die Wahrnehmung von *utility value,* jedoch in Abhängigkeit der situativen Kontrollwahrnehmung, gefunden werden (neben einem generellen, positiven Trend von Messzeitpunkt zu Messzeitpunkt). D.h., an Messzeitpunkten, an denen die situative Kontrollwahrnehmung höher ausgeprägt ist, fällt die Stärke des Effekts geringer aus. Im Subdatensatz „*Klassenführung*“ konnte ein negativer Treatmenteffekt auf die Wahrnehmung von *utility value* und *intrinsic value* beobachtet werden (neben einem generellen Abwärtstrend von Messzeitpunkt zu Messzeitpunkt). Mit jeder Wiederholung wird der negative Effekt allerdings schwächer. Kontrollkognitionen stellen zwar auch hier einen starken Prädiktor dar, moderieren aber nicht die Wirksamkeit bzw. die Stärke des Treatments.

Die Ergebnisse zu den Effekten des Treatments auf das situative emotionale Erleben – mediierte über die Valenzkognitionen – sowie die Interpretation der Bedeutsamkeit der Ergebnisse werden bis zur Tagung vorliegen.

26. Strategien zur Regulation von Motivation und Emotion in Lernkontexten – Entwicklung einer integrativen Taxonomie

Kristina Stockinger, Ulrike E. Nett, Markus Dresel
University of Augsburg, Germany

Theoretischer Hintergrund. Motivationale und emotionale Prozesse können das Wohlbefinden und akademische Leistungen von Lernenden bedeutend beeinflussen (Kriegbaum et al., 2018; Pekrun et al., 2022). Entsprechend stellen beide Konstrukte zentrale Stellschrauben selbstgesteuerten Lernens und damit einhergehender Regulationsprozesse dar (Boekaerts, 1999; Kim et al., 2020; Panadero, 2017). Aufgrund der konzeptuellen Gemeinsamkeiten zwischen Motivation und Emotion (vgl. Meyer &

Turner, 2006; Pekrun, in press) wurde in der Forschung bereits vereinzelt auf potenzielle Überschneidungen im Umgang mit motivationalen und emotionalen Problemen in Lernsituationen hingewiesen (z.B. Webster & Hadwin, 2015). Systematische Gegenüberstellungen fehlen jedoch bislang. Ziel dieses Beitrags ist es, ausgehend von etablierten pädagogisch-psychologischen Modellen der Motivationsregulation (z.B. Schwinger et al., 2007; Steuer et al., 2019) und Emotionsregulation (z.B. Harley et al., 2019; Pekrun & Perry, 2014) einer solchen Gegenüberstellung zu unterziehen und dadurch Gemeinsamkeiten und Unterschiede herauszuarbeiten, die die Grundlage für eine integrative Taxonomie bilden.

Fragestellung. In einem ersten Schritt wurde der Frage nachgegangen, welche Gemeinsamkeiten und Unterschiede zwischen Motivations- und Emotionsregulation auf theoretischer Ebene, insbesondere in Hinblick auf gängige Regulationsstrategien, existieren. Auf dieser Basis wurde eine integrative Strategietaxonomie entworfen und einem zweiten Schritt untersucht, inwiefern sich das Regulationsverhalten Studierender für motivationale versus emotionale Probleme unterscheidet, um die vorgeschlagene Taxonomie empirisch zu untermauern.

Methode. Strategiekataloge lernbezogener Motivationsregulation (Schwinger et al., 2007; Steuer et al., 2019; Wolters, 2003) und Emotionsregulation (Burić et al., 2016; Harley et al., 2019; Loderer & Pekrun, 2019) wurden anhand inhaltlicher Überschneidungen und Differenzen in eine gemeinsame Systematik überführt. Für eine erste empirischen Prüfung der postulierten Gemeinsamkeiten wurde anschließend eine Vignettenstudie durchgeführt. Hierbei berichteten 1.466 Studierende (M_{Alter}=22.8 J., *SD*=3.51; 70.5% weiblich) mehrerer deutscher/österreichischer Universitäten, inwiefern sie die in der integrativen Taxonomie repräsentierten Motivationsregulations- bzw. Emotionsregulationsstrategien in vier beschriebenen Lernsituationen anwenden würden. Die Situationen wurden in einem 2x2 Innersubjekt-Design als erwartungsbezogenes motivationales Problem (geringe Selbstwirksamkeit) bzw. emotionales Problem (Angst), und als wertbezogenes motivationales (fehlende Relevanz) bzw. emotionales Problem (Langeweile) während einer fiktiven Prüfungsvorbereitung präsentiert.

Ergebnisse und ihre Bedeutung. Existierende Motivationsregulations- und Emotionsregulations-Strategiekataloge überschneiden sich v.a. in Bezug auf Strategien der kognitiven Neubewertung. Beispielsweise finden sich entsprechende Strategien zur Steigerung des Kontrollerlebens (Motivationsregulation: fähigkeitsbezogene Selbstinstruktion; Emotionsregulation: kontrollbezogenes Reappraisal). Abweichungen lassen sich auf theoretischer Ebene v.a. auf unterschiedliche Regulationsziele zurückführen, die in der Emotionsregulationsliteratur deutlich breiter konzipiert sind: Während die Motivationsregulation zum Ziel hat, Lernprozesse aufrechtzuerhalten bzw. zu optimieren (Schwinger et al., 2007; Steuer et al., 2019), kann Emotionsregulation nicht nur instrumentell zur Förderung lern-/leistungsbezogener Prozesse eingesetzt werden, sondern auch direkt zur Beeinflussung emotionalen Erlebens (z.B. zur Reduktion oder Vermeidung unangenehmer Emotionen; vgl. Harley et al., 2019). Folglich wird beispielsweise Prokrastination z.B. als Emotionsregulationsstrategie zur Reduktion negativer Emotionen wie Angst oder Langeweile durch Situationsselektion betrachtet, nicht aber als Motivationsregulationsstrategie, deren Fokus ultimativ auf Lernprozessoptimierung liegt. Ferner zeigt sich empirisch anhand latenter Mittelwerte und

Korrelationen, dass der Strategieeinsatz eher nach Problemtyp (erwartungs-/kontroll- vs. wertbezogen) als nach Framing (motivationales vs. emotionales Problem) variiert. Diese Befunde stehen im Einklang mit der Annahme der Situationsspezifität von Motivationsregulation (vgl. Steuer et al., 2019). Zudem legen die Ergebnisse nahe, dass Regulationsstrategien, wie auch in unserer integrierten Strategietaxonomie angenommen, multifinal verwendet werden und unterstreichen die Angemessenheit sowie Bedeutung integrativer Theoriebildung für zukünftige Forschung und Praxis (z.B. bezüglich der Entwicklung kombinierter Fördermaßnahmen zur Regulation von Motivation und Emotion). Nächste Schritte zur weiteren Prüfung der integrativen Taxonomie in realen Lernkontexten sollen dahingehend diskutiert und eine Einbettung dieser Forschung in eine gesamtheitliche Perspektive auf selbstreguliertes Lernen reflektiert werden.

27. Through the lens of Scarcity. How perceived scarcity shapes student mindsets and learning experience?

Jessica Schulz, Matthias Nückles
Universität Freiburg, Germany

Much research has shown that to succeed academically students need to have a stable self-concept (Marsh, 1990) as well as self-efficacy (Schwarzer/Jerusalem, 1995) to regulate their learning processes and successfully formulate and follow their goals. Nevertheless, despite all the available knowledge about and diverse teachings on self-regulated learning, some students still fall behind. Are they simply less capable or is there something else that is hindering them? Is it possible that all these core elements of learning are being impeded by one overarching barrier?

For Carol S. Dweck this barrier is the students' mindset which plays a key role in motivation. According to the mindset theory students have either a fixed (entity theory of intelligence) or a growth mindset (incremental theory of intelligence). This determines whether they believe their intelligence is fixed and can't be changed or that, with motivation and effort, their intelligence and capabilities can be transformed. (Dweck, 2000) This study aims to go further, asking what students need to implement a growth mindset or what hinders them from developing one? We need to ask about the resources students actually have. We need to ask about scarcity. Following the definition from S. Mullainathan and E. Shafir scarcity is defined as a "subjective sense of having more needs than resources" (Mullainthan & Shafir, 2013: 86) or, in other words the feeling of "having less than you feel you need" (ibid.: 4). In several experimental studies they have shown that the core mechanism of scarcity is that it captures the mind in the form of a shortage of mental bandwidth. By this they mean that the mind focusses on whatever resource is scarce, capturing cognitive capacities as attention, executive control and even fluid intelligence.

Using a behavioral economic notion to explore students learning motivation, this study aims to investigate the perceived scarcity of students in their learning environment, bearing in mind the psychological consequences that scarcity possibly have for motivation, self-regulated learning processes and especially the question whether it correlates with an entity theory of intelligence. While there are several studies about the effect of induced scarcity on cognitive functioning (see: Shah, Mullainathan, Shafir: 2012) studies about the effect of scarcity in real-life settings is a scarce resource in itself.

Participants will be recruited from different universities and fields of studies. They will be asked to complete an online questionnaire. Measures for financial scarcity is a reduced version of the Psychological Inventory of Financial Scarcity (Van Dijk et al., 2021, Hilbert, Noordewier, van Dijk, 2022) and for time scarcity a reduced version of Van der Lippe's adjusted version of Garhammer's index of time pressure (Djupegot et al., 2017). Mindset will be investigated with the Implicit Theory of intelligence and the entity theory of intelligence questionnaire (Dweck, 2000). Following the study: *Money in the mental lives of the poor* (Shah, Zhao, Mullainathan, Shafir, 2018), participants will be presented with varoius scenarios and asked to indicate to what extent they think about several items, of which two are cost related and another two are money related. The aim is to get information about the core element of scarcity - to capture the mind. Related theories such as the self-concept (Marsh, 1990; Möller, Köller: 2004), self-efficacy (Scharzer/Jerusalem, 1995) and Basic Psychological Needs (Deci/Ryan, 2017) will also be taken into account in the questionnaire.

Since this study is still in progress at the time of submission, the hypotheses and expected results are described below.

- Higher scores of students' perceived scarcity correlates with lower basic psychological need satisfaction as well as lower self-efficacy scores.

- Perceived scarcity correlates with an Entity Theory of Intelligence

28. Validierung einer deutschen Adaptation der Brief Regulation of Motivation Scale

Alexander Luft, Olena Kryshko, Detlev Leutner
University Duisburg-Essen, Germany

Motivationale Selbstregulation als ein komplexer, mehr oder weniger bewusster Steuerungsprozess der eigenen Motivation umfasst drei Komponenten: Metamotivationales Wissen, metamotivationales Selbst-Monitoring und strategische Motivationskontrolle (Miele & Scholer, 2018). Es existieren zahlreiche Belege für den Einfluss motivationaler Selbstregulation auf das Lernverhalten und die akademische Leistung (Eckerlein et al., 2019; Kryshko et al., 2020; Schwinger & Stiensmeier-Pelster, 2012), wobei meistens die Motivationskontrolle anhand der Nutzung von Motivationsregulationsstrategien erfasst wurde. Messverfahren der Strategienutzung (Schwinger et al., 2007; Wolters & Benzon, 2013) sind jedoch lang, erfassen einzelne Motivationsregulationsstrategien, obwohl diese auch ineffektiv sein können, und erfragen nicht konkrete Motivationsprobleme. Als Alternative haben Kim et al. (2018) die „Brief Regulation of Motivation Scale" (BRoMS) vorgestellt, welche die generelle Tendenz zur motivationalen Selbstregulation in acht Items erfassen soll, sich zunehmend international etabliert (Bäulke et al., 2019; Garn & Morin 2021), allerdings noch nicht hinreichend validiert wurde. Ziel dieser Studie ist daher die Validierung einer deutschen Adaptation der BRoMS.

Die Skala wurde mit der „Back-Translation"-Methode (Brislin, 1970) ins Deutsche adaptiert und die Verständlichkeit der Items im Rahmen eines Interviews mit $N = 5$ Teilnehmenden überprüft. Die Validierung erfolgte unter Beachtung der von der AERA (2014) vorgeschlagenen Evidenzarten der Validität. Für die „Content-Oriented Evidence" wurden die Items von den Autor:innen hinsichtlich ihrer inhaltlichen Passung analysiert. Für die „Evidence Regarding Cognitive Processes" wurde geprüft, ob

die Interviewteilnehmenden die Items als Abfragen zur motivationalen Selbstregulation verstehen. Für weitere Evidenzen wurden die mithilfe einer Online-Studie erhobenen Selbstberichtsdaten von *N* = 350 Studierenden statistisch ausgewertet. Dabei wurden eine CFA zur Prüfung der erwarteten einfaktoriellen Skalenstruktur sowie Reliabilitätsanalysen durchgeführt („Evidence Regarding Internal Structure“). Außerdem wurden Korrelationen der BRoMS mit zwei weiteren Konstrukten der motivationalen Selbstregulation (motivationsregulatorische Selbstwirksamkeit und Strategienutzung) berechnet („Evidence Regarding Relationships With Conceptually Related Constructs“). Mit einer Methode zum Vergleich von Korrelationen (Hittner et al., 2003) wurde untersucht, ob die Korrelationen mit diesen Konstrukten höher sind als Korrelationen der BRoMS zu unverwandten Konstrukten (Verträglichkeit, Extraversion). Weiterhin wurde eine höhere Korrelation mit der Nutzung von Motivationsregulationsstrategien als mit motivationsregulatorischer Selbstwirksamkeit erwartet, weil die BRoMS eine Alternative zu Messverfahren des ersten Konstrukts sein sollte. Für die „Evidence Regarding Relationships With Criteria“ wurden theoriegeleitet Zusammenhänge mit Studienmotivationskomponenten (Erwartung, Wert, Kosten), Prokrastination, Anstrengung und akademischer Leistung (in Form der Durchschnittsnote) untersucht. Schließlich wurden für die „Evidence Based on Consequences of Tests” Messinvarianzen zwischen Gruppen hinsichtlich der Variablen Geschlecht, Migrationshintergrund und Bildungsherkunft überprüft.

Die Inhaltsanalyse zeigte, dass fünf Items die Strategienutzung zum Inhalt haben, während drei Items eher auf die Messung der motivationsregulatorischen Selbstwirksamkeit abzielen. Die Interviewteilnehmenden haben die Items als Fragen zur motivationalen Selbstregulation verstanden. Die CFA bestätigte die einfaktorielle Skalenstruktur (CFI = 0.95; TLI = 0.93; SRMR = 0.05; RMSEA = 0.08). Zudem ist die Skala hoch intern konsistent (α/ω >.80). Die BRoMS korreliert höher mit den beiden Konstrukten der motivationalen Selbstregulation als mit Verträglichkeit (r = 0.12; p = .039) oder Extraversion (r = 0.12; p = .042). Jedoch korreliert sie höher mit der motivationsregulatorischen Selbstwirksamkeit (r = 0.83; p < .001) als mit der Nutzung von Motivationsregulationsstrategien (r = 0.61; p < .001). Die BRoMS zeigt alle erwarteten Zusammenhänge mit Studienmotivationskomponenten (Erwartung: r = 0.30; p < .001; Wert: r = 0.26; p < .001; Kosten: r = -0.25; p < .001), Prokrastination (r = -0.37; p < .001), Anstrengung (r = 0.66; p < .001) und akademischer Leistung (r = -0.30; p < .001) auf. Schließlich ergeben sich vollständige Messinvarianzen hinsichtlich der Variablen Geschlecht und Bildungsherkunft sowie metrische Messinvarianz hinsichtlich des Migrationshintergrunds.

Insgesamt eignet sich die BRoMS mehr zur Erfassung motivationaler Selbstregulation als allgemeines Konstrukt, und weniger als Ersatz für die Messung der Strategienutzung. Weitere Implikationen bezüglich der zukünftigen Skalenverwendung werden diskutiert.

29. Motivationale Selbstregulation im Kontext Schule und Hochschule - Ein systematisches Literaturreview

Franzisca Hebald, Olena Kryshko, Detlev Leutner
Universität Duisburg-Essen, Deutschland

Motivationale Selbstregulation beschreibt die aktive Steuerung der eigenen Motivation im Rahmen des selbstregulierten Lernens (Wolters, 2011). Diese erfolgt über die Nutzung verschiedener

Motivationsregulationsstrategien, metamotivationales Selbst-Monitoring und metamotivationales Wissen (Miele & Scholer, 2018). Es gibt zahlreiche Studien über die positiven Effekte gelungener Motivationsregulation auf Lernprozesse von Schüler:innen und Studierenden (Daumiller & Dresel, 2018; Grunschel et al., 2016; Kryshko et al., 2022; Schwinger & Stiensmeier-Pelster, 2012). Bisher existiert allerdings keine Übersicht über das Forschungsfeld in diesem Themengebiet, das aus pädagogisch-psychologischer Sicht immer relevanter wird und sich zunehmend ausdifferenziert. Eine systematische Analyse der bisherigen Studien soll einen Überblick über den aktuellen Forschungsstand verschaffen und handlungsorientierte Implikationen für weitere Forschung erzielen.

Mit der Methode eines systematischen Literaturreviews (Grant & Booth, 2009; Moher et al., 2009) wurde die Forschungslandschaft im Bereich motivationaler Selbstregulation im Zeitraum der letzten 23 Jahre analysiert (ab der ersten Systematisierung unterschiedlicher Motivationsregulationsstrategien von Wolters, 1998). Dabei wurden folgende Forschungsfragen beantwortet: 1. Welche Aspekte motivationaler Selbstregulation standen bisher im Fokus der Forschung? 2. Mit welchen anderen Konstrukten wurde Motivationsregulation in Verbindung gebracht? 3. Welche Limitationen treten typischerweise bei Studien zu Motivationsregulation auf?

Die Literaturrecherche (vom 01.09. bis 01.12.2021) über die Datenbanken pedocs, PubPsych und Google Scholar beschränkte sich auf englisch- und deutschsprachige Studien sowie Schüler:innen und Studierende als Stichprobe. Es wurde beispielsweise mit folgenden Suchtermini gearbeitet: «Motivationsregulation», «motivationale Selbstregulation» oder «motivational self-regulation». Außerdem wurde via Google Scholar eingesehen, in welchen Publikationen Wolters (1998, 1999) zitiert wird, und die Literaturverzeichnisse passender Studien wurden nach thematisch ähnlichen Studien durchsucht (Schneeballprinzip). Durch die systematische Datenbankrecherche konnten zunächst *N* = 4860 Studien ausfindig gemacht werden. Mit der PRISMA-Flow-Diagramm-Methode (Stovold et al., 2014) wurden im nächsten Schritt kriteriengeleitet Studien ausgewählt. Exkludiert wurden u. a. Studien mit unpassender Stichprobe, Bachelor- und Masterarbeiten oder Duplikate. Nach einem mehrstufigen Screening-Prozess von Titel, Abstract und Volltext ergab sich eine finale Stichprobe von *N* = 96 Studien. Neben der Extrahierung allgemeiner Studiencharakteristika, wie z. B. Studiendesign oder Stichprobengröße, wurden für die Beantwortung von Forschungsfrage 1 folgende Aspekte von Motivationsregulation anhand des Modells von Miele & Scholer (2018) unterschieden: Strategienutzung, metamotivationales Wissen und metamotivationales Selbst-Monitoring. Außerdem wurden Kategorien für weitere untersuchte Konstrukte sowie Studienlimitationen gebildet, z. B. akademischer Erfolg oder Persönlichkeitsaspekte (Forschungsfrage 2) und Selbstberichtsverfahren oder korrelative Daten (Forschungsfrage 3).

Die Datenauswertung zeigte, dass deutlich mehr Studien im Kontext Hochschule (*N* = 79) als im Kontext Schule (*N* = 17) zu verzeichnen sind. Insgesamt liegt dabei der Fokus auf der Erforschung der Strategienutzung, insbesondere konkreter Strategiearten, die in *N* = 87 Studien erfasst wurden (Forschungsfrage 1). Die beiden anderen Aspekte motivationaler Selbstregulation, metamotivationales Wissen (*N* = 15) und Selbst-Monitoring (*N* = 0), wurden bisher selten bzw. gar nicht untersucht. Diese Forschungslücken können u. a. dadurch erklärt werden, dass hier teilweise valide Messinstrumente fehlen, an denen in Zukunft gearbeitet werden sollte. Ausgewählte

Ergebnisse zu Forschungsfrage 2 zeigen deutliche Themenschwerpunkte, die zusammen mit Motivationsregulation untersucht wurden: Selbstreguliertes Lernen (*N* = 56), akademischer Erfolg (*N* = 49) und motivationale Variablen (*N* = 34). Aufgrund der Vielzahl an Studien erscheint eine metaanalytische Ergebnissynthese zu den Zusammenhängen zwischen Motivationsregulation und diesen Konstrukten in Zukunft sinnvoll. Es werden außerdem weitere pädagogisch-psychologisch relevante, aber noch nicht ausreichend erforschte Themen in Verbindung mit Motivationsregulation aufgezeigt. Die drei meistgenannten Studienlimitationen (Forschungsfrage 3) sind Selbstberichtsverfahren (*N* = 30), eine nicht repräsentative Stichprobe (*N* = 32) und korrelative Daten (*N* = 15). In zukünftigen Forschungsvorhaben sollte daher beispielsweise der Einsatz multi-methodaler Verfahren oder experimenteller Studiendesigns stärker berücksichtigt werden.

Im Rahmen des Posterbeitrags werden diese und weitere Ergebnisse des Literaturreviews detailliert präsentiert sowie Lösungsansätze bezüglich gefundener Problematiken diskutiert. Der bereitgestellte Forschungsüberblick kann dabei als Vorbereitung für eine Metaanalyse dienen.

30. Zum Einfluss von Feedback auf Lernerfolg und affektive Schüler*innenmerkmale

Svenja Boegel, Mathias Ropohl
Universität Duisburg Essen, Deutschland

Theoretischer Hintergrund

Speziell für den naturwissenschaftlichen Unterricht stellt Feedback eine essentielle Lernunterstützung dar. Schüler*innen zeigen in Experimentiersituationen, insbesondere bei der Anwendung der Variablenkontrollstrategie, Lernschwierigkeiten, welche angepasste Lernprozesse erfordern (Scheuermann, 2017).

Die Schüler*innen müssen in einem experimentellen Lernszenario die Informationen der Lernmaterialien zum Experimentieren und die Feedbackinformationen verarbeiten. Daraus entsteht eine komplexe Lernsituation, welche die Schüler*innen kognitiv und affektiv überfordern kann (Leutner et al., 2014). In empirischen Untersuchungen wurde analysiert, inwiefern Feedback, den Aufbau von Kompetenzen beim Experimentieren unterstützt. Wollenschläger et al. (2011) konnten herausstellen, dass Feedback, welches die Kompetenzen der Schüler*innen anhand von Kriterien adressiert, wirksam ist. Hild et al. (2020) konnten herausfinden, dass Schüler*innen von Feedback zum aktuellen Leistungsstand profitieren. Ropohl und Scheuermann (2018) konnten zeigen, dass Feedback, welches Informationen zum Lernziel, zum Lernstand und zu den nächsten Schritten enthält zu einem besseren Abschneiden der Schüler*innen führt. Die Ergebnisse aller drei Studien lassen vermuten, dass affektive und kognitive Konstrukte einen Einfluss auf die Wirkung des Feedbacks haben. Mögliche Einflussfaktoren sind die Motivation der Schüler*innen sowie ihr wahrgenommener Cognitive Load beim Experimentieren (Ryan & Deci, 2000; Sweller, 1988). Den Einfluss dieser affektiven und kognitiven Merkmale gilt es beim feedbackgestützten Experimentieren hinsichtlich Lernbereitschaft und Lernerfolg vertieft zu untersuchen.

Forschungsfragen

Forschungsfrage 1: Welches Feedback schätzen Schüler*innen als unterstützend und motivierend ein?

Forschungsfrage 2: Inwiefern wird der Effekt von Feedback auf den Lernzuwachs fachmethodischer Kompetenzen beim Experimentieren unter Anwendung der Variablenkontrollstrategie über die Motivation und den wahrgenommenen Cognitive Load vermittelt?

Methode

Um dasjenige Feedback zu identifizieren, das aus Schüler*innenperspektive besonders unterstützend und motivierend wirkt (Forschungsfrage 1), ist eine Querschnittserhebung (Herbst 2022) geplant. Die geplante Stichprobe umfasst vier Klassen eines Gymnasiums in NRW der Jahrgangsstufe 8 (N ~ 90). Die Schüler*innen erhalten zu einem von ihnen geplanten Experiment individuelles Feedback und dürfen für die weitere Arbeit eine von drei Feedbackvarianten auswählen. Diese Auswahl muss begründet und das Feedback bewertet werden. Das Fachwissen und die kognitiven Grundfähigkeiten werden als Kontrollvariablen erhoben. Das aus diesen Daten abgeleitete Feedback findet in der Interventionsstudie Einsatz.

Mit Blick auf Forschungsfrage 2 ist eine Interventionsstudie geplant (Herbst 2023). Das aus den Daten der Querschnittserhebung als motivierend identifizierte Feedback wird als unabhängige Variable in seiner Ausgestaltung variiert. Der Lernzuwachs naturwissenschaftlich-experimenteller Kompetenzen wird als abhängige Variable erfasst. Des Weiteren werden das Fachwissen und die kognitiven Grundfähigkeiten als Kontrollvariablen erhoben. Die Motivation und der Cognitive Load fließen als Mediatorvariablen ein. In der Interventionsgruppe 1 erhalten die Schüler*innen ein motivierendes Feedback. In Interventionsgruppe 2 eine Feedbackform, die zusätzlich den Cognitive Load reduziert. Die Kontrollgruppe bearbeitet anstelle des Feedbacks eine vergleichbare unterrichtliche Aktivität.

Ergebnisse und ihre Bedeutung

Die Varianzanalyse mit Messwiederholung aus einer Voruntersuchung (Ropohl & Scheuermann; 2018) belegt, dass Schüler*innen im Rahmen der Intervention in Abhängigkeit von den erhaltenen Feedbackinformationen einen unterschiedlich starken Lernzuwachs im Planen von Experimenten aufweisen ($F_{Prä\text{-}Post}$ (2, 190) = 11.989, p < .001, η^2 = .112). Mit Hilfe eines Bonferroni Post-hoc-Tests für den Prä- und Post-Messzeitpunkt wurde ermittelt, dass Feedback, welches Informationen zum Lernziel, zum Lernstand und zu den nächsten Schritten enthält, den höchsten Lernzuwachs bewirkt (p < .001, η^2 = .44). Basierend darauf enthält das Feedback in der unter Kapitel 3 erwähnten Querschnittserhebung immer diese drei Informationen. Die Lernwirksamkeit dieser Feedbackinformationen soll repliziert und mit den von Schüler*innen als motivierend empfundenen Feedbackelementen kombiniert werden. Im Rahmen der geplanten Untersuchung soll analysiert werden, ob die Schüler*innen qualitativ hochwertigere Experimente planen, wenn sie Feedback erhalten, dass sie als motivierend und unterstützend einschätzen. Die konkreten Ergebnisse der Querschnittserhebung werden auf der GEBF 2023 präsentiert werden.

31. Wer profitiert motivational? Effekte von Leistungsfeedback in einem Test zur Multiple Document Comprehension und Über- bzw. Unterschätzung auf die Feedbackwahrnehmung und Motivation sich zu verbessern

Theresa Zink[1,3], Carolin Hahnel[2,5], Ulf Kroehne[2], Tobias Deribo[2], Nina Mahlow[3], Cordula Artelt[3], Frank Goldhammer[2,5], Johannes Naumann[4], Cornelia Schoor[3]

[1]Friedrich-Alexander-Universität Erlangen-Nürnberg; [2]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; [3]Leibniz-Institut für Bildungsverläufe; [4]Bergische Universität Wuppertal; [5]Zentrum für internationale Bildungsvergleichsstudien (ZIB)

Multiple Document Comprehension (MDC) ist eine wichtige Kompetenz für Studierende, da sie meist mehrere Dokumente zu einem Thema lesen und verstehen müssen (Schoor, Hahnel, Mahlow et al., 2020). Bisherige Studien zur Förderung von MDC-Kompetenz beschäftigten sich zwar mit Effekten von Feedback auf die Motivation bezüglich der Anwendung bestimmter Lesestrategien (z.B. Maier & Richter, 2014), allerdings gibt es noch wenig Forschung dazu, wie sich Feedback oder eine Über- bzw. Unterschätzung der eigenen Kompetenzen auf die Motivation, sich in MDC zu verbessern, auswirken. Wir nehmen an, dass beide Aspekte wichtige Prädiktoren der Motivation, sich in MDC zu verbessern, sind. Im Allgemeinen kann positives Feedback die Motivation fördern, wohingegen negatives Feedback die Motivation verringern kann (z.B. Bandura, 1986; Zimmermann, 2000). Gleichzeitig kann eine moderate Überschätzung der eigenen Kompetenz im Gegensatz zu einer extremen Über- bzw. Unterschätzung positive Effekte auf die Motivation haben (z.B. Baumeister et al., 1989; Lee, 2021).

Das Ziel unserer Studie ist es herauszufinden, ob die im Feedback rückgemeldete MDC-Leistung sowie ihre Über- bzw. Unterschätzung einen Effekt auf die Motivation, sich in MDC zu verbessern, hat. Darüber hinaus wird untersucht, ob die Feedbackwahrnehmung die Motivation, sich in MDC zu verbessern, beeinflusst und gleichzeitig von der rückgemeldeten Leistung bzw. einer Über- bzw. Unterschätzung vorhergesagt werden kann.

Im Rahmen dieser Studie nahmen N = 285 Studierende (72% weiblich, M_{Alter} = 25.6, SD_{Alter} = 4.8) von drei deutschen Universitäten an einem standardisierten MDC-Test (Schoor, Hahnel, Artelt et al., 2020) teil. Nach der Bearbeitung des Tests erhielten die Studierenden ein individuelles Feedback. Zusätzlich wurden sie gebeten, Stellung zu ihrem MDC-bezogenen Selbstkonzept, ihrer Feedbackwahrnehmung (Nützlichkeit, Akzeptanz und Spezifität) sowie zu ihrer Motivation, sich in MDC zu verbessern, zu beziehen. Die Motivation, sich zu verbessern, wurde nach dem Erwartungs-Wert-Modell (Eccles & Wigfield, 2020) in Form von drei Facetten abgebildet: Selbstwirksamkeitserwartung (r=.55), Nützlichkeit (ω=.90) und eingeschätzte Kosten (ω=.66). Die Über- bzw. Unterschätzung wurde mittels standardisierter Residualscores ermittelt, indem das MDC-bezogene Selbstkonzept (ω=.92) auf die ermittelte Leistung im MDC-Test regrediert und standardisiert wurde. Zusätzlich wurde als weitere Variable ein quadratischer Term des Residualscores gebildet um Ausprägungen der Über- bzw. Unterschätzung (z.B. moderate Überschätzung) zu untersuchen. Die Auswertung der Daten erfolgt mittels mehrerer latenter Regressionsmodelle.

Die Ergebnisse zeigten teilweise statistisch-signifikante Effekte der rückgemeldeten Leistung auf die Motivation, sich in MDC zu verbessern. Je negativer die rückgemeldete Leistung im Feedback war, desto höher waren die wahrgenommenen Kosten einer Verbesserung (ß=-.22; p=.01). Es fand sich kein statistisch-signifikanter Zusammenhang zwischen dem (quadrierten) Residualscore und der Motivation, sich in MDC zu verbessern. Ähnliche Ergebnisse zeigten sich auch für die Feedbackwahrnehmung: Die Feedbackakzeptanz war bei einer positiv rückgemeldeten Leistung höher als bei einer negativen Rückmeldung (ß=.26; p<.001) und es gab keine Effekte des

(quadrierten) Residualscores auf die Feedbackwahrnehmung. Im Hinblick auf die Effekte der Feedbackwahrnehmung auf die Motivation, sich zu verbessern, zeigte sich, dass je nützlicher das Feedback wahrgenommen wurde, desto ausgeprägter war die Selbstwirksamkeitserwartung (ß=.57; *p*<.001) sowie die wahrgenommene Nützlichkeit einer MDC-Verbesserung (ß=.78; *p*<.001). Mediationseffekte der Feedbackwahrnehmung auf die Motivation, sich in MDC zu verbessern, wurden weder für die rückgemeldete Leistung noch die Residualscores gefunden.

Die Ergebnisse weisen darauf hin, dass eine positiv rückgemeldete Leistung mit höherer Feedbackakzeptanz und geringeren wahrgenommenen Kosten, sich in MDC zu verbessern, einhergeht. Das bedeutet, dass in der vorliegenden Studie die rückgemeldete Leistung relevanter für die Feedbackwahrnehmung und Motivation, sich in MDC zu verbessern, ist als eine bestimmte Über- bzw. Unterschätzung der eigenen Kompetenz (z.B. moderate Überschätzung). Zusätzlich deuten die Analysen darauf hin, dass eine hohe wahrgenommene Feedbacknützlichkeit relevant für die Motivation, sich in MDC zu verbessern, ist. Zusammenfassend profitieren motivational vor allem Studierende, die bereits eine hohe MDC-Kompetenz besitzen, sowie diejenigen, die das Feedback als nützlich wahrnehmen.

32. Folgen der Förderung studentischer Forschung an der TU Dresden. Eine Befragung von Teilnehmenden am Programm FOSTER

Anne Jaschan, Franziska Schulze-Stocker
TU Dresden, Deutschland

Im Zuge der Diskussionen um Effekte der Exzellenzförderung auf die Gruppe der Studierenden gewinnt die Förderung der Forschungsorientierung bzw. des Forschenden Lernens im Studienverlauf zunehmend an Relevanz (Reinmann, 2016). U. a. wird Forschendes Lernen als „Instrument zur Förderung forschungsbezogener Kompetenzen“ (Lübcke et al., 2017, S. 192) verstanden und als vielversprechend bzgl. des Verhältnisses von Hochschulbildung und Arbeitsmarkt eingeschätzt (Wissenschaftsrat, 2015). Forschungsorientierte Lern- und Lehrangebote bzw. die gezielte Förderung studentischer Forschung stellen somit ein Segment frühzeitiger Nachwuchsförderung und der Karriereentwicklung dar.

Das Poster stellt die Konzeption und erste Ergebnisse der Begleitforschung des Programms FOSTER (Funds for Student Research) vor, welches als Teil der Exzellenzstrategie „TUD 2028 – Synergy and beyond“ studentische Forschung strukturiert fördert und sichtbar macht. Als primäre Zielgruppe werden Studierende v. a. in ihren Forschungsvorhaben, bei der Ausrichtung von studentischen Veranstaltungen mit Forschungsbezug, der Teilnahme an Forschungswettbewerben sowie der Präsentation eigener Forschungsprojekte bei Konferenzen/Tagungen oder durch Publikationen finanziell unterstützt. Zudem erhalten sie passgenaue Weiterbildungsangebote. Als sekundäre Zielgruppe zählen Lehrende bzw. Nachwuchswissenschaftlicher/innen, welche an der TU Dresden Lehr-Lern-Formate entwickeln und durchführen wollen, die Forschendes Lernen fokussieren und damit den Forschungsbezug aktiv in der Lehre herstellen. Durch FOSTER sollen damit vor allem *research-based learning* und *research-based teaching* als aktives Lernen und Lehren gefördert werden (Healey & Jenkins, 2009).

Im Rahmen der Begleitforschung wird eine quantitative Befragung der Teilnehmenden mit offenen und geschlossenen Fragen genutzt, um den Zugewinn der Förderung für die studentische Forschung zu ergründen. Die offenen Fragen werden dabei inhaltsanalytisch ausgewertet (Schreier, 2012; 2014). Das verwendete Kategoriensystem unterscheidet zwischen wissenschaftlichen Interessen (bezogen auf Inhalte, Methodik und Kontakte/Netzwerke), Forschungsmotivation, wissenschaftlicher Werdegang, Veröffentlichungen und Folgeprojekten. Als reflexive Komponente dient die Frage nach dem wissenschaftlichen Werdegang, da sie sowohl im Antragsformular als auch im Abschlussbericht beantwortet werden muss.

Die anschließende Analyse des Datenmaterials fokussiert die Folgen der Förderung hinsichtlich der Festigung wissenschaftlicher Interessen und der Präsentation bzw. Veröffentlichung von Forschungsergebnissen sowie hinsichtlich des Werdegangs.

33. Studienabbruchtypologien im Fernstudium

Nathalie Krüger, David Wick, Birgitt Erdwien, Tobias Filusch
Europäische Fernhochschule Hamburg GmbH, Deutschland

Die Erforschung des Phänomens Studienabbruch ist in der Hochschul- und Bildungslandschaft bereits seit vielen Jahren von großem Interesse. In der Studienabbruchforschung existieren verschiedene theoretische Konzepte zur Erklärung von Studienabbrüchen, welche vorwiegend psychologische, kulturelle und soziale Rahmenbedingungen berücksichtigen (vgl. Heublein 2017; Sarcletti & Müller, 2011). Fernhochschulen als eigenständige Hochschulform finden allerdings in diesen Studien nur selten Berücksichtigung. Dabei gewinnt das (private) Fernstudium immer mehr an Bedeutung, welches gerade innerhalb der Pandemie seit 2020 einen deutlichen Aufwind erlebt (Thuy & Höllermann, 2011; forsa, 2021). Das spezifische Studierendenklientel an Fernhochschulen – v.a. berufsbegleitend Studierende, ältere Studierende oder Studierende ohne klassische Hochschulzugangsberechtigung (nicht-traditionelle Studierende) – dürfte jedoch abweichende Bedingungsfaktoren mit sich bringen. Die hier präsentierte Untersuchung beleuchtet, welche Einflussfaktoren besonders relevant für Studienabbrüche an privaten Fernhochschulen sind und inwiefern sich daraus verschiedene Typen von Studienabbrecher*innen identifizieren lassen. Sie ist Teil des Projekts Studienerfolge und -abbrüche im Fernstudium (SaFe), welches durch das BMBF gefördert wird. Es zielt darauf ab, das Forschungsdesiderat bei der Betrachtung von Studienerfolg und Studienabbruch an Fernhochschulen zu schließen und zu überprüfen, welche Bedingungsfaktoren für Studienerfolge und Studienabbrüche im Fernstudium von besonderer Relevanz sind. Das Herzstück des Forschungsvorhabens liegt in der hier vorgestellten Längsschnittstudie, mit der die Identifizierung von Studienabbruchstypen realisiert wird. Dies bildet den Ausgangspunkt für die Entwicklung und Implementierung eines Online-Self-Assessments, welches Studierenden Hinweise auf mögliche Entwicklungspotentiale sowie individuell maßgeschneiderte Unterstützungsangebote liefert.

Bei der Längsschnittstudie handelt es sich um eine Fragebogenerhebung mit zwei Erhebungszeitpunkten im Abstand von sechs Monaten, innerhalb derer aktive Fernstudierende der Europäischen Fernhochschule Hamburg und der APOLLON Hochschule für Gesundheitswirtschaft, Bremen, zu ihren individuellen Lebens- und Lernbedingungen, zu kritischen Momenten in ihrem

Studium sowie zu einer etwaigen Studienabbruchneigung befragt werden. Der Fragebogen basiert auf den Ergebnissen einer im Rahmen des SaFe-Projektes ebenfalls durchgeführten Interviewstudie mit Alumni, Kohortenbesten und Abbrecher*innen der Europäischen Fernhochschule Hamburg, einer Analyse hochschulinterner Studierendendaten sowie eines umfangreichen Literaturstudiums. Zur Berechnung der Typologien wird ein probabilistischer Ansatz (Hybridmodell) gewählt. Dieser reagiert besonders sensitiv auf Wirkungsgefüge höherer Ordnung, was die Auffindung qualitativ unterschiedlicher Subpopulationen im Kontext von Studienabbruchsprozessen fördert (vgl. Martens, 2000). Dazu werden die erhobenen Daten teils skalenanalytisch mithilfe des Rasch-Modells und teils gruppenanalytisch mittels der Latent Class Analyse analysiert. Anschließend werden nach Maßgabe der Studienabbruchsneigung und der diese bedingenden Faktoren mit Hilfe von Hybridmodellen Studierendentypologien bestimmt und ihre Profile expliziert (Martens, Erdwien & Ramm, 2008; Martens, 2000).

Die Längsschnittstudie befindet sich derzeit in der ersten Erhebungsphase und wird nach ersten Ergebnissen im November 2022 im Mai 2023 mit der zweiten Erhebung abgeschlossen. Demnach können zum Zeitpunkt der Tagung bereits erste Ergebnisse zur Typologienbildung vorgestellt werden.

34. Validierung eines Fragebogens zur Erfassung statistikspezifischer epistemologischen Überzeugungen mittels psychometrischer Netzwerke – Ergebnisse einer Pilotstudie

Eric Klopp[1], Leopold Hamminger[2], Robin Stark[1]
[1]Universität des Saarlandes, Deutschland; [2]Europa Virtuelle Volkshochschule

Theoretischer Hintergrund

Epistemologische Überzeugungen sind ein wichtiger Prädiktor wissenschaftlichen Denkens in einer Disziplin (Fischer et al., 2014), wobei allgemeine und domänenspezifische Überzeugungen unterschieden werden können (Muis et al., 2006). Insbesondere domänenspezifische Überzeugungen besitzen für wissenschaftliches Denken eine besondere Relevanz. Die vorliegende Pilotstudie beschreibt die Validierung des auf Domäne Statistik adaptierten EBI-AM Fragebogens (Peter et al., 2016), der statistikspezifischen Absolutismus und Multiplizismus erfassen soll, mittels psychometrischer Netzwerke (Constatini et al., 2014).

Ein Netzwerk besteht auf Knoten, welche die Variablen (Items) repräsentieren, die durch Kanten verbunden sind. Die Kanten repräsentieren paarweise Interaktionen zwischen den Knoten, die im Sinne von Partialkorrelationen zu verstehen sind. Ein wichtiges Element bei der Konstruktion eines psychometrischen Netzwerkes ist die Bestimmung der Kanten, die letztendlich im Netzwerk verbleiben. Hierzu werden üblicherweise Regularisierungsverfahren eingesetzt (Epskamp et al., 2017). Weiterhin lassen sich in Netzwerken Communities nachweisen; das sind Gruppen von Knoten, die untereinander hohe Gemeinsamkeiten aufweisen. Somit sollten die zu einer Skala gehörenden Items als Community im Netzwerk nachzuweisen sein (Briganti et al., 2018; Golino & Epskamp, 2017).

Rosman et al. (2017) zeigten, dass bei Informatikstudenten sich längsschnittlich eine Zunahme domänenspezifischen Absolutismus und eine Abnahme domänenspezifischen Multiplizismus zeigt. Die Autoren begründen dies in der spezifischen Wissensstruktur der Domäne Informatik, die überwiegend aus mathematischen Inhalten besteht. Mathematischen Wissen wird durch die Stabilität

seiner Inhalte charakterisiert, sind diese einmal bewiesen, ändern diese sich nicht mehr. Daher sollte der Erwerb mathematischen Wissen zu hohem Absolutismus und niedrigen Multiplizismus führen. Ähnliches dürfte auch somit für andere mathematische Disziplinen, wie sie auch die Statistik darstellt zu erwarten sein. Insbesondere dürfte in einer querschnittlichen Betrachtung ein positiver Zusammenhang zwischen Absolutismus und ein negativer Zusammenhang zwischen Multiplizismus und Domänenwissen zu erwarten sein.

Fragestellung

Im Folgenden werden die Hypothesen untersucht, dass sich a) ein Netzwerk zeigt, dessen Knoten sich in zwei Communities gruppieren, welche die jeweiligen Items zur Erfassung von statistikspezifischem Absolutismus und Multiplizismus enthalten und b) sich ein positiver Zusammenhang zwischen statistikspezifischen Absolutismus bzw. negativer Zusammenhang zwischen statistikspezifischen Multiplizismus zeigt.

Methode

Das psychometrische Netzwerk wurde anhand einer Stichprobe von 262 Studierenden geschätzt. Statistikspezifischer Absolutismus (α=.78, 12 Items) und Multiplizismus (α=.74, 11 Items) wurden reliabel erfasst. Zusätzlich wurde selbsteingeschätztes, statistisches Wissen (α=.94) sowie die Abiturmathematiknote erfasst.

Das Netzwerk wurde durch eine regularisierte Schätzung mittels EBICglasso (λ=.25; Epskamp et al., 2017) ermittelt. Zur Absicherung des Netzwerkstruktur wurde weiterhin die Akkuratheit der Kanten sowie die Stabilität des Netzwerkes mithilfe der Bootstrap-Methode mit 2000 Bootstrapstichproben untersucht (Epskamp et al., 2018). Zur Bestimmung der Community-Struktur wurde der Spinglass-Algorithmus eingesetzt (Yang et al., 2016).

Ergebnisse

Das Netzwerk wies eine gute Modellpassung auf ($\chi^2(151)=162.15$, p=.25, CFI=0.97, RMSEA=0.030, SRMR=0.038). Hinsichtlich der Stabilität zeigte sich nur eine geringe Abweichung der geschätzten Werte vom Mittelwert des Bootstraps (max. Abweichung von Schätzer und Booststrapmittelwert: 0.025). Auch wies das Netzwerk eine ausreichende Stabilität auf (CS=.28). Der Spinglass-Algorithmus ergab zwei Communities, die jeweils denn die den beiden Skalen zugeordneten Items enthielt. Die Absolutismus- und Multiplizismusskala korrelierten nicht signifikant (r=-.07, p=.270). Hinsichtlich der Zusammenhänge zwischen epistemologischen Überzeugungen und Statistikwissen zeigte sich lediglich eine knapp signifikante, negative Korrelation zwischen Multiplizismus und Statistikwissen (r=-.13, p=.050).

Die Ergebnisse zeigen, dass sich die angenommen Struktur des auf die Domäne Statistik übertragenen EBI-AM mittels eines psychometrischen Netzwerks nachweisen lässt, was somit für die Validität des adaptierten Fragebogens spricht. Hinsichtlich der Korrelation der Skalen zeigte sich lediglich ein signifikanter, hypothesenkonformer Zusammenhang von Multiplizismus mit Statistikwissen, der allerdings nur ein kleine Effektstärke zeigt. Dieser Befunde spricht für die konvergente Validität des adaptierten Instruments. Wahrscheinlich ist die geringe Effektstärke sowie der fehlende Zusammenhang zwischen Absolutismus und Statistikwissen darauf zurückzuführen,

dass lediglich selbsteingeschätztes statt objektiv vorhandenes Statistikwissen, erfasst wurde. Weiterführende Studien sollten somit objektiv Statistikwissen erfassen.

35. Strategien nach Fehlern in Simulationen - Effekte von Vorwissen

Nicole Heitzmann[1], Matthias Stadler[1], Constanze Richters[1], Anika Radkowitsch[2], Ralf Schmidmaier[3], Marc Weidenbusch[4], Martin R. Fischer[4]

[1]Lehrstuhl für Empirische Pädagogik & Pädagogische Psychologie, LMU München; [2]IPN Leibniz-Institut für die Didaktik der Naturwissenschaften und der Mathematik Kiel; [3]Medizinische Klinik und Poliklinik IV, LMU Klinikum; [4]Institut für Didaktik und Ausbildungsforschung in der Medizin, LMU Klinikum

Theoretischer Hintergrund

Der Erwerb von diagnostischen Kompetenzen ist ein wichtiges Ziel der akademischen Ausbildung in unterschiedlichen Fächern wie der Medizin oder der Lehramtsausbildung (Heitzmann et al., 2019). Für das Diagnostizieren sind diagnostische Aktivitäten wie das Formulieren von Hypothesen und das Generieren von Evidenz zentral (Heitzmann et al., 2019). Simulationen sind für den Erwerb von komplexen Fähigkeiten wie den diagnostischen Kompetenzen für Lernende mit unterschiedlichem Vorwissen hoch effizient (Chernikova et al., 2020). Welche Mechanismen das Lernen mit Simulationen so effektiv machen, ist jedoch bislang nicht umfassend geklärt. Eine Möglichkeit ist, dass Lernende in Simulationen, anders als in realen Handlungssituationen, andere Möglichkeiten haben auf gemachte Fehler zu reagieren. Mögliche Strategien nach einem Fehler sind beispielsweise die Generierung weiterer Evidenz innerhalb der Simulation um Hypothesen zu bestätigen oder zu verwerfen. Eine weitere Strategie besteht in der Reflexion der bisherigen Handlungssequenz ohne das direkt auf den Fehler reagiert werden muss. Lernende mit unterschiedlichem Vorwissen verwenden vermutlich unterschiedliche Strategien (Mayr et al., 2011). Unklar ist weiter, inwiefern die Strategien stabil sind.

Fragestellungen

FF1: Wie unterscheiden sich Strategien nach einem Fehler in einer Simulation von Lernenden mit unterschiedlichem Vorwissen?

FF2: Sind die genutzten Strategien nach einem Fehler in einer Simulation stabil?

Methode

In einer medizinischen Simulation diagnostizierten Teilnehmer:innen fiktive, realistische, Patientenfälle, indem sie eine elektronische Patientenakten begutachteten und radiologische Befunde anforderten (siehe Radkowitsch et al., 2022). Zur Beantwortung von FF1 wurde ein quasi-experimentelles einfaktorielles Design (*N*=98) mit Lernenden auf drei Vorwissensniveaus verwendet (niedrig-mittel-hoch). Zur Beantwortung der FF2 wurden die Kontrollgruppen aus zwei weiteren Interventionsstudien reanalysiert (*N*=89). Anschließend wurden alle Fälle ausgeschlossen, in denen kein Fehler protokolliert wurde (*N*=87, mit n=295 Beobachtungen). Die Beobachtungen wurden gemäß drei Anpassungsstrategien (Reflexion, Evidenzgenerierung oder beides) kodiert. Anschließend wurden die Sequenzen mit dem von Gabadinho et al. (2011) vorgeschlagenen Tracemining-Algorithmus analysiert.

Ergebnisse

FF1: Lernende mit hohem Vorwissen nutzten mit gleicher Wahrscheinlichkeit Evidenzgenerierung (48,1%) und Reflexion. Sowohl Lernende mit mittlerem (70,1%) als auch mit niedrigem Vorwissen (66,1%) nutzten häufiger Evidenzgenerierung. Lernende mit hohem Vorwissen nutzten mehr Zeit, bevor sie nach einem Fehler fortfuhren. Dies war unabhängig davon, ob weitere Evidenz generiert ($F_{2,179}$=4.24; p=.016; η^2=.045) oder reflektiert wurde ($F_{2,139}$=2.88; p=.030; η^2=.040). Ein logistisches Regressionsmodell in dem Evidenzgenerierung und Vorwissen die Wahrscheinlichkeit einer korrekten Diagnose vorhersagten, zeigte eine signifikante Interaktion zwischen den Prädiktoren. Es zeigte sich ein signifikanter Haupteffekt der Evidenzgenerierung (b=-.99; p=.013; OR=0.37) sowie signifikante Unterschiede zwischen Lernenden mit mittlerem und hohem Vorwissen, und zwischen mit niedrigem und hohem Vorwissen. Die Wahrscheinlichkeit eine korrekte Diagnose zu stellen, verringerte sich mit zunehmender Evidenzgenerierung bei Lernenden mit hohem und niedrigem Vorwissen, während sie bei Lernenden mit mittlerem Vorwissen anstieg.

FF2: Es konnten drei Cluster identifiziert werden: (1) Reflektierende Lernende, (2) Evidenzgenerierende Lernende und (3) variable Lernende, die beide Strategien verwendeten.

Diskussion

Dieselbe Strategie nach einem Fehler kann für bestimmte Lernende hilfreich sein, aber negative Folgen für Lernende mit anderen Lernvoraussetzungen haben. Lernende mit mittlerem Vorwissen gehen hypothesengeleitet, deduktiv beim Diagnostizieren vor (Rikers et al., 2000). Möglicherweise lag dementsprechend die Fehlerursache nur bei Lernenden mit mittlerem Vorwissen in einer unzureichenden Evidenzgenerierung. Bei Lernenden mit hohem Vorwissen scheint es darum sinnvoll eine instruktionale Unterstützung in der Simulation zu ergänzen, die keine weitere Evidenzgenerierung zulässt, sondern sie zur weiteren Reflexion ihres bisherigen Vorgehens anregt. Lernende mit geringem Vorwissen waren nicht in der Lage, von weiteren Evidenzen oder von Reflexion zu profitieren. Anstatt aussagekräftige Beweise zu generieren, griffen sie wahrscheinlich auf eine Versuch-und-Irrtum-Strategie zurück. Möglicherweise hätten sie z. B. von einer elaborierten Form des Feedbacks oder von einem Lösungsbeispiel profitiert. Die Ergebnisse vorliegender Studie können im Kontext von Fehlern in Simulationen Ansatzpunkte zur Personalisierung von Unterstützungsmaßnahmen für Lernende mit unterschiedlichem Vorwissen bieten.

36. Wahrnehmung und Anwendung widersprüchlicher Daten im Licht von alternativen Theorien – Eine Eyetracking Studie

Sabine Meister, Anna Beniermann, Annette Upmeier zu Belzen
Humboldt-Universität zu Berlin, Germany

Im Prozess der naturwissenschaftlichen Erkenntnisgewinnung werden Daten generiert und genutzt, um bestehendes Wissen zu erweitern. Dabei können initiale Theorien durch Daten geändert oder verworfen und durch neue Theorien ersetzt werden (Chinn & Brewer, 1993). Im Bereich der Naturwissenschaftsdidaktik wird dieses Vorgehen als theoretischer Rahmen zur Beschreibung kognitiver Lernprozesse in Form der *Conceptual Change Theorie* genutzt. Innerhalb der *Conceptual Change Theorie* werden Lernprozesse durch einen kognitiven Konflikt ausgelöst, der mit der initialen Theorie nicht erklärbar ist und nach einer alternativen und plausiblen Theorie verlangt (Posner et al., 1982). Angelehnt an naturwissenschaftliche Prozesse, hat sich die Präsentation von

widersprüchlichen Daten (*anomalous data*) zur Auslösung eines kognitiven Konflikts zum Lernen naturwissenschaftlicher Theorien in Forschung und Lehr-Lern-Prozessen durchgesetzt (Chinn & Brewer, 1998, Limón, 2001). Jedoch zeigen bisherige Studien, dass der Einsatz widersprüchlicher Daten nur selten zur Änderung von initialen Theorien führt (Potvin, 2021). Dabei wird der Wahrnehmung der Widersprüchlichkeit ein hoher Stellenwert zugeschrieben, welche den Lernprozess mit Daten erst ermöglicht (Meister & Upmeier zu Belzen, 2021). Ohlsson (2013) und Potvin (2021) plädieren dafür, Theorieänderungen eher im Sinne einer *cognitive utility* zu verstehen, bei der es darum geht, dass nicht Daten selbst zur Theorieänderung führen, sondern die Nützlichkeit der jeweiligen Theorie bei ihrer Anwendung, um diese Daten zu erklären. Dies wird als Übergang von einer wahrnehmungsbasierten zu einer anwendungsbasierten Perspektive von *Conceptual-Change*-Forschung beschrieben (Ohlsson, 2013). Darüber hinaus gibt es mehrere Studien, die zeigen, dass widersprüchliche Daten effektiver in Lernprozessen genutzt werden können, wenn bereits eine alternative und plausible Theorie für die Lernenden vorliegt (Hemmerich et al., 2016; Potvin, 2021). Somit stellt sich die Frage nach dem Stellenwert widersprüchlicher Daten im Lernprozess mit Blick auf die Beziehung zwischen Datenwahrnehmung und Anwendung von alternativen Theorien.

Fragestellung

In dieser Studie untersuchen wir, inwiefern die Datenwahrnehmung und die Anwendung von initialer gegenüber alternativer Theorie mit der Stellung von widersprüchlichen Daten im Lernprozess zusammenhängt. Folgende Forschungsfragen werden dabei adressiert.

Inwiefern werden widersprüchliche Daten im Vergleich zu unterstützen Daten ohne und mit vorheriger Bereitstellung alternativer Theorien unterschiedlich visuell wahrgenommen?

Inwiefern führt die Wahrnehmung widersprüchlicher Daten ohne und mit vorheriger Bereitstellung alternativer Theorien zur Theorieänderung?

Methode

Mit einem Testinstrument wurden initiale Theorien zum Thema Populationsdynamik von Biologielehramtsstudierenden (N=20) erfasst (Meister & Upmeier zu Belzen, 2021). Anschließend wurden den Proband*innen widersprüchliche und bestätigende Daten zu ihrer initialen Theorie in Form von Liniendiagrammen auf einem Bildschirm gezeigt, der an ein Eyetracking-Gerät angeschlossen war. Die aufgenommenen Blickbewegungen wurden anhand der Metriken Fixationsanzahl und durchschnittliche Fixationsdauer analysiert (Holmqvist et al., 2011). Theorieänderung nach Präsentation der widersprüchlichen Daten wurde per Testinstrument erhoben.

In einer anschließenden Studie werden den Studierenden alternative Theorien zur Erklärung von Populationsdynamiken vorgelegt, bevor ihnen die zu ihrer initialen Theorie widersprüchlichen Daten gezeigt werden, wobei erneut Blickbewegungen gemessen werden. Die Theorien der Studierenden werden ebenfalls vor und nach der Konfrontation mit den Daten per Testinstrument erhoben.

Ergebnisse und ihre Bedeutung

Ergebnisse aus der der ersten Erhebung ohne Bereitstellung einer alternativen Theorie zeigen, dass widersprüchliche Daten häufiger und länger wahrgenommen werden als bestätigende Daten.

Dennoch erfolgt meist keine Änderung der initialen Theorie, was vorherige Ergebnisse aus *Conceptual-Change*-Studien stützt (Potvin, 2021). Jedoch zeigt sich, dass Wissen über alternative Theorien vorhanden ist, diese aber als unwahrscheinlicher charakterisiert werden. Dies lässt vermuten, dass die vorherige Bereitstellung alternativer Theorien, die Anwendung dieser bei der Erklärung widersprüchlicher Daten fördern könnte. Dazu werden die Ergebnisse der Folgeerhebung herangezogen, welche auf der Tagung vorgestellt werden.

37. Einstellungen und Argumente zu Controversial Science Issues

Anna Beniermann, Annette Upmeier zu Belzen
Humboldt-Universität zu Berlin, Germany

Theoretischer Hintergrund

Gesellschaftliche Diskurse über wissenschaftliche Themen verdeutlichen Diskrepanzen zwischen wissenschaftlichem Konsens und öffentlicher Meinung (Hornsey & Fielding, 2017), aus denen gesellschaftliche Kontroversen erwachsen können (van der Linden, 2019). Gesellschaftlich kontrovers diskutierte wissenschaftliche Themen, über die wissenschaftlicher Konsens besteht, werden als *Controversial Science Issues* (CSI; Beniermann et al., 2021) bezeichnet. Solche Themen werden als Kontexte im pädagogischen Ansatz *Socioscientific Issues* (SSI; Sadler, 2004) genutzt und gewinnen im Zuge veränderter Anforderungen an naturwissenschaftlichen Unterricht an Bedeutung (Upmeier zu Belzen & Beniermann, 2020).

Um sich eine Meinung zu diesen Themen zu bilden, bedarf es Argumentationskompetenz (Lazarou & Erduran, 2021), die mit individuellen Einstellungen zum Gegenstand (Means & Voss, 1996) sowie mit Wissen über *Nature of Science* (NOS; Khishfe, 2021) zusammenhängt. Auch Vertrauen in Wissenschaft beeinflusst, wie wissenschaftliche Erkenntnisse interpretiert werden (Sinatra et al., 2014) und damit die Fähigkeit zum Argumentieren. Unsicherheit, z. B. im Sinne von fehlendem Wissen oder fehlendem Konsensus (Gustavson & Rice, 2020) hängt mit einem Vertrauen in Wissenschaft zusammen (Chinn et al., 2018). Dies macht Einstellungen zu CSI, NOS-Wissen sowie Vertrauen in Wissenschaft zu wichtigen Aspekten für die Untersuchung von Argumenten zu CSI.

Die meisten Studien zur Klassifikation von Argumenten fokussieren auf deren Struktur auf der Grundlage von *Toulmin's Argument Pattern* (TAP; Toulmin, 1958), während wenige Studien den Inhalt von Argumenten (z. B. Jafari & Meisert, 2021) oder Begründungen als Teil von Argumenten (z. B. Lobato & Zimmerman, 2019) untersuchten.

Fragestellungen

Inwieweit hängen NOS-Wissen und Vertrauen in Wissenschaft mit der Einstellung zu CSI zusammen?

Wie werden Einstellungen zu CSI begründet?

Methode

Biologie-Lehramtsstudierende (N=44) wurden mit einem Online-Fragebogen zu Einstellungen zu vier CSI (Evolution: Beniermann, 2019; Klimawandel: Christensen & Knezek, 2015; GMF: Rose et al., 2019; Impfungen: Betsch et al., 2018), NOS-Wissen (Urhahne et al., 2008) und Vertrauen in

Wissenschaft (Nadelson et al., 2014) befragt. Zur Erfassung von Begründungen zu CSI als Teil von Argumentationen wurden die Studierenden mit Aussagen zu diesen CSI konfrontiert, die jeweils einen wissenschaftlichen Konsens abbilden (Beniermann et al., 2021). Zustimmungswerte wurden auf einer Rating-Skala erfasst und im offenen Antwortformat begründet sowie Gründe für eine potentielle Einstellungsänderung angegeben. Antworten wurden im Rahmen einer qualitativen Inhaltsanalyse (Mayring, 2010) gemäß TAP (Toulmin, 1958) mit einem Kategoriensystem zu CSI (Beniermann et al., 2021) analysiert. Begründungen von Einstellungen werden exemplarisch für GMF dargestellt.

Ergebnisse

Höheres NOS-Wissen hing positiv mit Einstellungen zu Evolution (r=.248; p<.05) und Impfungen (r=.270; p<.05) zusammen. Vertrauen in Wissenschaft korrelierte positiv mit Einstellungen zu Evolution (r = .353; p < .01), Klimawandel (r=.218; p< 05) und Impfungen (r=.423; p<.01). Einstellungen zu GMF hingen nicht mit NOS-Wissen oder Vertrauen in Wissenschaft zusammen. Die Studierenden stimmten überwiegend dem wissenschaftlichen Konsens in Bezug auf Klimawandel, Evolution und Impfungen zu, während die Mehrheit der Studierenden bzgl. GMF unsicher war (n=20) oder den wissenschaftlichen Konsens ablehnte (n=10).

Während Begründungen positiver Einstellungen zu GMF sich überwiegend auf wissenschaftliche Kontrollmechanismen oder Analogien bezogen, wurden negative Einstellungen zu GMF mehrheitlich subjektiv begründet (z. B. Emotionen). Studierende mit unsicherer Einstellung zu GMF begründeten meistens mit Wissenslücken oder vermuteter fehlender Evidenz. Als Gründe für potentielle Einstellungsänderungen wurden überwiegend zusätzliche wissenschaftliche Studien oder die Aneignung von Fachwissen angegeben.

Schlussfolgerungen

Die Ergebnisse deuten auf Kontextabhängigkeit des Zusammenhangs von NOS-Wissen und Vertrauen in Wissenschaft mit Einstellungen zu CSI hin. Obwohl die quantitativen Daten keinen Zusammenhang zwischen NOS-Wissen und Vertrauen in Wissenschaft und Einstellungen zu GMF nahelegen, verweisen die qualitativen Daten auf die Bedeutung dieser beiden Faktoren, wenn Studierende positive Einstellungen mit Vertrauen in Kontrollmechanismen sowie potentielle Einstellungsänderungen mit zusätzlichen wissenschaftlichen Studien begründen. Die Ergebnisse lassen vermuten, dass Unsicherheit in Bezug auf wissenschaftliche Evidenz sowie individuelle Wissenslücken die Bildung von Einstellungen zu CSI behindern und affektive Faktoren an Bedeutung gewinnen.

38. Service-Learning als innovatives Lehr-Lern-Format zur Förderung der Third Mission von Hochschulen – Ergebnisse einer Umfrage an deutschen Hochschulen

Katrin Arendt, Lisa Stark, Anja Friedrich, Roland Brünken, Robin Stark
Universität des Saarlandes, Deutschland

Beim Service-Learning wird ein universitärer Kurs (Seminar, Vorlesung etc.) mit einer Dienstleistung der Studierenden (Praxistätigkeit, ehrenamtliches Engagement etc.) kombiniert (z.B. Hofer, 2019). Das Konzept ist v.a. im englischsprachigen Raum verbreitet; positive Wirkungen hochqualitativer Service-Learning-Angebote u.a. auf die persönliche und akademische Entwicklung Studierender

sowie die Förderung gewinnbringender Interaktionen der Universitäten mit Kooperationspartner*innen aus ihrem Umfeld werden berichtet (Billig et al., 2005; Eyler, 2011). Damit scheinen Service-Learning-Angebote zur Verbindung zwischen Hochschule und Gesellschaft (Altenschmidt & Hagemus-Becker, 2019), zur Stärkung des Vertrauens in die Nützlichkeit (Hofer, 2019) und somit zur Third Mission von Hochschulen beizutragen. In der Forschung für den deutschsprachigen Raum müssen diese Befunde noch weiter überprüft werden. Um Aussagen über die Qualität von Service-Learning-Veranstaltungen treffen zu können, ist es notwendig, aufbauend auf Qualitätskriterien systematisch Instrumente zur Erfassung der Qualität von Service-Learning-Veranstaltungen zu entwickeln (Sonnberger & Leitzmann, 2020).

Das Ziel der vorliegenden Umfrage bestand erstens darin, eine Bestandsaufnahme von Service-Learning-Angeboten an deutschen Hochschulen und ihren Rahmenbedingungen aus Dozierendensicht durchzuführen und zweitens, ein Instrument zur Erfassung der Qualität von Service-Learning-Veranstaltungen (QuaSeL) zu entwickeln und zu pilotieren.

An der Umfrage nahmen *N*=122 Dozierende (Alter: *M*=42.85, *SD*=11.14; 50,42% weiblich; 47,54% Professor*innen, 18,03% promovierte, 21,31% nicht-promovierte wissenschaftliche Mitarbeiter*innen) neun deutscher Hochschulen teil. Sie beantworteten sechs geschlossene (z.B.„Wie schätzen Sie die Kosten für Service-Learning-Veranstaltungen im Vergleich zu anderen Lehrveranstaltungen? niedriger--höher“) und vier offene Fragen (z.B.„Was erschwert die Durchführung von Service-Learning-Veranstaltungen aus Ihrer Sicht?“) zu Aufwand, Kosten und Nutzen sowie Gründen für und gegen Service-Learning. Die Antworten der offenen Fragen wurden von zwei unabhängigen Ratern kodiert (Mayring, 2010). Zudem beantworteten diejenigen, die angaben, Service-Learning-Veranstaltungen durchgeführt zu haben, den neu konzipierten QuaSeL-Fragebogen. Für zehn theoretisch abgeleitete Qualitätsmerkmale von Service-Learning (gesellschaftlicher Bedarf, definierte Ziele, Kooperation der Beteiligten, Lernen in fremden Lebenswelten, Bestandteil des Studiums, Kompetenzerwerb der Studierenden, Reflexion, Begleitung der Studierenden, Evaluation und Qualitätsentwicklung, Anerkennung und Würdigung) wurden in einem kooperativen Austauschprozess auf Basis der Definitionen (Sonnberger & Leitzmann, 2020) insgesamt 42 Items abgeleitet und nach einem Expertenurteil optimiert. Diese bewerteten die Dozierenden auf einer 4-Punkte-Skala (1=trifft nicht zu bis 4=trifft zu) bezüglich ihrer Service-Learning-Veranstaltung.

Bei der Bestandsaufnahme zeigte sich, dass 33,94% der befragten Dozierenden Service-Learning vor der Umfrage kannten. Von diesen ordneten *n*=17 (45,95%) mindestens eine ihrer Lehrveranstaltungen als Service-Learning ein. Neben dem Arbeitsaufwand (*M*=3.30, *SD*=0.59) und den Kosten (*M*=2.97, *SD*=0.82) für die Durchführung von Service-Learning wurde auch der Nutzen von Service-Learning für die Studierenden (*M*=3.37, *SD*=0.74) von den Dozierenden im Vergleich zu anderen Lehr-Lern-Formaten als eher hoch eingeschätzt. Häufig genannte Gründe für die Durchführung von Service-Learning waren der Kompetenzerwerb der Studierenden (4 Nennungen) sowie die Theorie-Praxis-Verbindung (3 Nennungen); Gründe dagegen waren die vermeintlich fehlende Eignung curricularer Inhalte für Service-Learning (30 Nennungen) sowie der höhere Aufwand (18 Nennungen). Als Schwierigkeiten wurden Probleme beim Finden und Arbeiten mit Kooperationspartner*innen genannt (17 Nennungen) sowie die fehlende Etablierung des Konzepts an

der eigenen Hochschule (11 Nennungen). Dozierende wünschten sich für Service-Learning-Veranstaltungen v.a. mehr Informations- und Weiterbildungsangebote (21 Nennungen) sowie finanzielle und personelle Unterstützung (11 Nennungen). Dozierende, die Ihre Veranstaltung als Service-Learning einordneten, schätzten deren Qualität eher hoch ein (Gesamtskala: M=3.31, SD=0.54). Trotz der kleinen Stichprobengröße sind die Reliabilitätswerte für alle QuaSeL-Subskalen (.64<α<.84; Kubinger, 2009) außer für die Skala zum Kompetenzerwerb (α=.47) als akzeptabel bis sehr gut einzuordnen; acht Items wurden aufgrund zu geringer Trennschärfen aus der Analyse entfernt und für eine optimierte Fragebogenversion revidiert.

Die Ergebnisse zeigen auf, dass Service-Learning in Deutschland nicht weit verbreitet ist, Maßnahmen wie Dozierenden-Workshops deren Einsatz jedoch verstärken könnten. Der explorative Charakter und die geringe, nicht-repräsentative Stichprobe, insbesondere bezogen auf die Erprobung des QuaSeL-Fragebogens, stellen Limitationen dieser Studie dar. Der Fragebogen sollte überarbeitet, an einer repräsentativen Stichprobe validiert und um weitere Einschätzungsperspektiven (Studierende, Kooperationspartner*innen) ergänzt werden.

39. Was wissen Studierende über KI? Entwicklung und Validierung eines Tests zur Messung von AI Literacy

Marie Hornberger, Arne Bewersdorff, Claudia Nerdel
Professur für Fachdidaktik Life Sciences, TUM School of Social Sciences and Technology, Technische Universität München

Künstliche Intelligenz (KI) ist durch Sprachassistenten und weitere Tools in unserem Alltag angekommen und durchdringt alle Lebensbereiche (Paaß & Hecker, 2021). Um den gesellschaftlichen Wandel durch KI demokratisch zu gestalten, sollen KI-Kompetenzen in der breiten Bevölkerung aufgebaut werden (Deutscher Bundestag, 2020). Die Hochschulen können zum Aufbau dieser KI-Kompetenzen beitragen, indem sie technische und nicht-technische KI-Spezialist*innen ausbilden (Deutscher Bundestag, 2020; Witt et al., 2020).

Stand der KI-Kompetenzen bei Studierenden

Für eine bedarfsorientierte Entwicklung eines entsprechenden Studienangebots zu KI ist es wichtig zu klären, welche Kompetenzen diese Zielgruppe erwerben soll und auf welchem Stand sich diese Kompetenzen bei Eintritt in einen Kurs oder ein Studienprogramm befinden (Witt et al., 2020). So kann im Sinne einer Eingangsdiagnostik das Vorwissen der Studierenden berücksichtigt und der Lernerfolg gezielt unterstützt werden (Köller et al., 2020). Der aktuelle Stand der KI-Kompetenzen bei Studierenden in Deutschland ist aufgrund des Mangels an einschlägigen empirischen Untersuchungen noch weitgehend unklar. Qualitative Studien deuten darauf hin, dass Lernende Fehlvorstellungen über KI haben (Lindner & Berges, 2020; Sulmont et al., 2019). Insbesondere Studierende nicht-technischer Fächer sollten Wissenslücken bezüglich KI aufweisen, da mathematische Fähigkeiten und Programmierkenntnisse als die beiden größten Hürden beim Lernen über KI identifiziert wurden (Sulmont et al., 2019).

Forschungsfrage: Über welche KI-Kompetenzen verfügen Studierende in unterschiedlichen Fächern?

AI Literacy

Zur Klärung der Forschungsfrage wird der Ist-Stand der *AI Literacy* von Studierenden erhoben. Unter diesem Konstrukt versteht man grundlegende Kompetenzen im Bereich KI, die für eine mündige Teilhabe in einer von KI-Technologien durchdrungenen Gesellschaft wichtig sind (Long & Magerko, 2020; Ng et al., 2021). Nach der Definition von Long und Magerko (2020) umfasst AI Literacy eine Reihe von Kompetenzen, die es einem Individuum ermöglichen, KI-Technologien kritisch zu bewerten, effektiv mit KI zu kommunizieren und zusammenzuarbeiten und KI als ein Werkzeug online, zuhause und am Arbeitsplatz zu nutzen (Long & Magerko, 2020). Auf Basis eines Literatur-Reviews schlagen Long und Magerko (2020) 17 verschiedene Kompetenzen vor (z.B. „Distinguish between technological artefacts that use and do not use AI."), die anhand von fünf Fragen gruppiert werden („What is AI?", „What can AI do?", „How does AI work?", „How should AI be used?", „How do people perceive AI?").

Testentwicklung

Da kein etabliertes Testinstrument zur Messung dieser Kompetenzen identifiziert werden konnte, wurde für diese Studie ein neues Testinstrument entwickelt. Hierzu wurden in einem ersten Schritt 57 Items theoriegeleitet zu den Kompetenzen von Long und Magerko (2020) entwickelt. Für die Itemgenerierung wurden außerdem etablierte KI-Lehrbücher (z.B. Ertel, 2016; Russell & Norvig, 2022) herangezogen. Für die Entwicklung geeigneter Distraktoren wurden Forschungsergebnisse über Fehlvorstellungen zu KI verwendet (z.B. Lindner & Berges, 2020; Sulmont et al., 2019). Die Items wurden von drei Fachwissenschaftler*innen im Feld der KI bezüglich ihrer fachlichen Korrektheit überprüft und ggf. überarbeitet. Außerdem wurde mit drei Studierenden die Methode des Lauten Denkens durchgeführt, um Verständnisprobleme zu identifizieren und ggf. zu beheben. In einer folgenden Pilotstudie mit 24 Teilnehmer*innen wurden deskriptive Maße für die psychometrische Qualität des Testinstruments erhoben. Die Items wurden anschließend erneut überarbeitet. Die finale Testversion wurde mithilfe einer deskripivstatistischen Itemanalyse (Itemschwierigkeit und -trennschärfe) zusammengestellt. Das finale Testinstrument umfasst 30 Multiple-Choice-Aufgaben und eine Sortieraufgabe.

Hauptstudie

Für die Haupterhebung wird eine Stichprobe von Studierenden (geplant N>300) unterschiedlicher Studiengänge an mehreren Hochschulen in Deutschland rekrutiert, um die Reliabilität und Validität des Tests zu überprüfen und Aussagen über den Kompetenzstand treffen zu können. Zur Überprüfung der Validität des Tests werden eine Selbsteinschätzung der eigenen Kompetenzen, Interesse an KI, Erfahrung mit KI und das Studienfach erhoben. Die Haupterhebung wird zu Beginn des Wintersemesters 2022/23 durchgeführt. Die Daten werden nach der Item-Response-Theorie ausgewertet. Die Testentwicklung sowie erste Ergebnisse sollen auf der Tagung zur Diskussion gestellt werden.

40. Innovativer Lernraum und seine Evaluation

Jochen Weißenrieder, Veronika Weber-Schopp
RWU (Hochschule Ravensburg-Weingarten)

„*Hochschulen sollten innovative Raum- und Ausstattungskonzepte über Pilotprojekte auf ihrem Campus umsetzen und erproben. Dadurch können sie Praxiserfahrung sammeln und die Annahme abstrakt-konzeptioneller Ideen von der tatsächlichen Zielgruppe überprüfen bzw. verbessern lassen.*“ (Wissenschaftsrat 2022)

Mit dem LearningLab der Hochschule Ravensburg-Weingarten (RWU) wird diese Forderung nach einem innovativen Lehr-Lern-Raum-Konzept umgesetzt. Unser Konferenzbeitrag behandelt das implementierte Konzept, einen ungewöhnlichen Evaluationsansatz sowie erste empirische Ergebnisse.

Die Gestaltung der Lehr-Lern-Räume in Hochschulen verändert sich nur langsam und spiegelt in Hörsälen und Seminarräumen noch eine Lehrendenzentrierte Lehre wieder (Reinmann-Rothmeier & Mandl, 2006). Die Lehre entwickelt sich jedoch weiter. Studierendenzentrierung gewinnt an Bedeutung, Konstruktivistische Lerntheorie manifestiert sich in Förderung und Forderung von aktivierender Lehre. Projektorientierte Lehre entspricht dabei sowohl der „Hands-On“ Tradition der Hochschulen für Angewandten Wissenschaften als auch dem Ruf nach praxisnahen Methoden zur Förderung von FutureSkills. Denn Hochschulen müssen sich auch verändern, weil die Arbeitswelt nicht nur im Zuge der Digitalisierung andere Absolvent:innen erwartet.

Welche Räume die Hochschulen hierfür brauchen, wird vielfach diskutiert. Die RWU hat daher wie andere Hochschulen auch ein Pilotprojekt gestartet. Das LearningLab an der RWU lädt ein, unterschiedliche Lehr-Lern-Settings auszuprobieren und zu kombinieren. Es besteht aus einem Veranstaltungsraum sowie angrenzenden Coworkingbereich. Der Coworkingbereich dient im Falle hoher Teilnehmerzahlen gleichzeitig als Erweiterung des LearningLabs, insbesondere für Gruppenarbeitsphasen. Die Möblierung stellt das Gegenkonzept zu Sitzreihen der traditionellen Räume dar, ist für agile und kreative Arbeit im Stehen angelegt und unterstützt den Austausch aller Anwesenden.

Die mediendidaktische Ausstattung zielt darauf ab, vorgegebene räumliche Grenzen didaktisch (und gedanklich) zu überwinden. Hierfür stehen **vier** vielfältige Lehr-Lern-Settings (**4ALL**) zur Verfügung, die miteinander kombiniert werden können. Es handelt sich um das analoge Setting, ein digitales Setting, ein Setting für immersives Lernen sowie eine offenes Setting für CoWorking und Werkstätten. Allen Settings ist gemein, dass sie den aktivierenden Charakter mit jeweiligen Arbeitsmaterialien unterstreichen, die für alle Anwesenden (**4ALL**) und nicht nur für die Lehrperson stets griffbereit stehen. Zudem hilft fachkundliches Personal bei der Auswahl (oder Entwicklung) der jeweils passenden Lernumgebung für die geplanten Veranstaltungen.

In der evaluativen Begleitung interessieren uns Aspekte der Akzeptanz, Usability und Nutzenbewertung: Auf quantitativer Ebene erfassen wir zunächst die reinen Nutzungszahlen über das eingesetzte Buchungssystem. Auf qualitativer Ebene erheben wir subjektive Bewertung der Nutzer:innen, die auch Verhaltensabsichten, Emotionen und Eindrücke umfassen.

Die Herausforderung besteht darin, dass der Aufwand für die Datenerhebung auf Seiten der Nutzer:innen dem Ertrag durch die z.T. eher kurze Nutzung angemessen sein muss. Daher setzen wir

sowohl im LearningLab als auch im Coworkingbereich auf kurze, flexible Formate, die sich dem hoch agilen Raum anpassen und den Nutzer:innen nicht zu viel abverlangen.

Im Anschluss an Veranstaltungen im LearningLab setzen wir sogenannte Feedback-Kärtchen (Postkarten) ein, mit der die Nutzer:innen abwechselnd nach einer allgemeinen Einschätzung gefragt werden und oder gezielt spezifische Eindrücke, Verhaltensabsichten, Emotionen o. Ä. erhoben werden. Ergänzt wird dies durch Leitfadengestützte Feedbackgespräche mit den Lehrenden. Da der Coworkingbereich noch heterogener und beweglicher ist, setzen wir hier eine Umfrage durch Kugelung bzw. Ballotage ein. Die „wöchentliche Evaluationsfrage“ gibt ein aktuelles Bild zum Bedarf Studierender und Lehrender. Für weitergehende Fragestellungen wird gegen Ende jedes Semesters eine Onlinebefragung der Nutzer:innen des LearningLabs und des Coworkingbereiches durchgeführt.

Erste Ergebnisse zeigen zwei deutliche Erfolge: Zum einen sind soziale Austauschformate und Gruppenarbeitsphasen deutlich einfacher umzusetzen, als in traditionellen Settings. Zum Anderen ermöglicht die flexible Ausstattung auch, von Austauschformen schnell in ruhigere, individuelle Arbeitsphase zu wechseln. Die technische Infrastruktur wird ebenfalls positiv aufgenommen, angefangen von der Verfügbarkeit von Steckdosen und W-Lan bis hin zu gemeinsamer Nutzung der großen Bildschirme.

41. Lernbezogene Effektivität von Virtual Reality im Bildungsbereich

Matthias Conrad, David Kablitz, Stephan Schumann
University of Konstanz, Germany

Zusammenfassung

Die Studie untersucht lernbezogene Effekte des Einsatzes von Virtual Reality (VR) im Kontext von Lehren und Lernen. Im Mittelpunkt steht die Frage nach der Vorteilhaftigkeit dieser Technologie im Vergleich zu anderen Medien. Zudem wird untersucht, inwieweit deklaratives und prozedurales Wissen durch VR gefördert wird und welche Art von Lernumgebung sich beim Einsatz von VR als besonders vorteilhaft erweist. Zur Beantwortung dieser Fragen wurde ein systematisches Review auf Basis der PRISMA-Methode nach Liberati et al. (2009) durchgeführt. Die Ergebnisse zeigen, dass sich der Einsatz von VR besonders für den Erwerb von prozeduralem Wissen eignet.

Theoretischer Hintergrund

Der Einsatz von VR ermöglicht die Realisierung komplexer Lehr- und Lernarrangements und kann auf diese Weise zur Erreichung spezifischer Lernziele beitragen. Es wird angenommen, dass diese Technologie dazu in der Lage ist, den Erwerb von Fachwissen und praktischen Fähigkeiten effektiv zu fördern (Radianti et al., 2020). Während "explorative VR-Umgebungen" vor allem dazu geeignet sind, eine virtuelle Umgebung selbstständig zu erkunden, zielen "VR-basierte Trainings- und Konstruktionsumgebungen" auf den Erwerb von handlungsbezogenen Fähigkeiten ab (Schwan & Buder, 2006). Trotz vielfältiger instruktionaler Gestaltungsmöglichkeiten existieren bislang nur wenige systematische Befunde in Bezug auf die Erreichung kognitiver Lernziele durch Einsatz von VR, so dass weiterführende Erkenntnisse hierüber wünschenswert sind, um diese Technologie wirksam in der Bildungspraxis einzusetzen.

Forschungsfragen

Im Rahmen der Studie werden folgende Forschungsfragen verfolgt:

a) Ist der Einsatz von VR in Bezug auf den Wissenserwerb effektiver als bei anderen Medien?

b) Fördert der Einsatz von VR den Erwerb von deklarativem oder prozeduralem Wissen?

c) Welche Art von Lernumgebung ist beim Einsatz von VR am effektivsten?

Forschungsmethode

Zur Beantwortung dieser Fragen wurde ein systematisches Review nach der PRISMA-Methode von Liberati et al. (2009) durchgeführt. Folgende Datenbanken wurden für die Recherche verwendet: APA PsycArticles, APA PsycInfo, ERIC und Teacher Reference Center. Auf diese Weise wurden 871 Artikel identifiziert, von denen nach einem mehrstufigen Selektionsprozess 30 relevante Studien übrigblieben, die eine Gesamtstichprobe von 2.384 Probanden umfassten.

Ergebnisse

Die identifizierten Studien liefert deutliche Hinweise auf die lernbezogene Überlegenheit von VR-basierten Lehr- und Lernsettings im Vergleich zum Einsatz von analogen Referenzmedien bzw. Unterrichtsformen (z.B. Lehrbuch, praktische Übung). Die Gegenüberstellung von Effekten zwischen dem Einsatz von VR und anderen elektronischen Medien (z.B. Video, Tablet-Computer, Augmented Reality) zeigt hingegen keine Überlegenheit in Bezug auf den Wissenserwerb. Ferner zeigt sich, dass sowohl der Erwerb von deklarativem als auch von prozeduralem Wissen vor allem in Trainings- und Konstruktionsumgebungen gefördert werden, bei denen das praktische Handeln im Mittelpunkt steht.

Diskussion und praktische Implikationen

Die Ergebnisse zeigen, dass der Einsatz von VR den Wissenserwerb effektiv unterstützen kann. Die Ergebnisse machen jedoch auch deutlich, dass der Einsatz von VR anderen Medien oder Unterrichtsmethoden per se nicht überlegen ist. Die Befunde weisen darauf hin, dass der Einsatz von VR vor allem für den Erwerb von prozeduralem Wissen geeignet ist, insbesondere wenn die zugrundeliegende Lernumgebung auf die Anwendung von praktischem Handeln abzielt.

Literatur

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine, 6*(7), Article e1000100, 1-28.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*, Article 103778, 1-29.

Schwan, S. & Buder, J. (2006). *Virtuelle Realität und e-learning*. Abrufbar unter https://www.e-teaching.org/didaktik/gestaltung/vr/vr.pdf (Abgerufen am 23.09.2022).

42. Der schmale Grat zwischen Publikationsverzerrung und Qualitätssicherung: Das Peer-Review-Verfahrens als geeignetes Einschlusskriterium für Systematic Reviews in der empirischen Bildungsforschung?!

Carolin Quenzer-Alfred, Daniel Mays
Universität Siegen, Germany

Systematic Reviews gewinnen zunehmend an Bedeutung und Popularität in der Bildungsforschung, um das Forschungsfeld systematisch zusammenzufassen und hinsichtlich einer Fragestellung zu synthesieren. Dabei handelt es sich um eine ursprünglich aus der Medizin stammende Methode zur

systematischen Aggregation von Forschungsdaten, deren Leitlinien häufig unverändert übernommen werden. So auch das Peer-Review-Verfahren als weit verbreitetes und akzeptiertes Einschlusskriterium von Forschungsbeiträgen. Durch unabhängige Fach- und MethodenexpertInnen dient es der Kontrolle und Sicherstellung einer möglichst hohen Qualität und Relevanz wissenschaftlicher Manuskripte und gilt als Markenzeichen einer erfolgreichen, hochwirksamen wissenschaftlichen Fachzeitschrift (Müller, 2008). Unter dem Gesichtspunkt der Publikationsverzerrung wird es aber auch in diesen Disziplinen kontrovers diskutiert – insbesondere im Hinblick auf statistische Verzerrungen mit Bevorzugung positiver bzw. signifikanter Ergebnisse, GutachterInnenverzerrung, sprachliche Verzerrung, Verzerrung durch das Fachgebiet oder geschlechtsspezifische Verzerrung (Armstrong, 1997; Goldbeck-Wood, 1999). Für den Bereich der Bildungswissenschaften eröffnet sich ein weiteres Problem der Leitlinienpassung: Nicht jede Forschungsarbeit wird als Zeitschriftenartikel veröffentlicht und nicht alle Zeitschriften verfügen über ein Peer-Review-Verfahren zur Qualitätskontrolle. So haben Kuhberg-Lasson, Singleton und Sondergeld (2014) gezeigt, dass zwischen 1995 und 2008 über 50 % der in der Erziehungswissenschaft generierten Outputs ohne Peer-Review veröffentlicht wurden. Die Schlüsselposition von Artikeln ohne Peer-Review in Bildungsstudien wird in einer Umfrage (n=643) von Bambey (2016) bestätigt. Obwohl Peer-Review von 93% der Umfrageteilnehmer als wichtig für das wissenschaftliche Arbeiten eingestuft wird, misst ein großer Teil der Umfrageteilnehmer (61%) auch Forschungsarbeiten ohne Peer-Review eine hohe Relevanz zu. Zusammen mit dem Problem der Publikationsverzerrung stellt sich die Frage, ob Peer-Review als prävalentes Kriterium ein berechtigtes und geeignetes Einschlusskriterium in systematischen Reviews ist, um den Forschungsstand in der Bildungsforschung umfassend zu erfassen, oder ob es für den Bereich der Erziehungswissenschaft andere Kriterien braucht, um die Community über den State of the Art umfassend zu informieren.

Präsentiert werden exemplarische Daten einer vergleichenden Studie zweier systematischer Übersichtsarbeiten von Quenzer-Alfred et al (2022) und Mertens et al. (2022) zum Einsatz digitaler Medien für SchülerInnen mit zusätzlichem oder sonderpädagogischem Förderbedarf im inklusiven, integrativen oder exkludierenden Unterricht publiziert mit und ohne Peer-Reviewverfahren. In einem interdisziplinären Forschungsteam wurden acht nationale und internationale Datenbanken gescreent und empirische Studien mit Peer-Reviewverfahren eingeschlossen. Das Verfahren wurde für ein zweites Review von Forschungsarbeiten ohne Peer-Reviewverfahren in zusätzlich auf graue Literatur spezialisierte Datenbanken mit identischen ReviewerInnen, Forschungsfragen und Suchstrings repliziert. Eine Qualitätsbewertung wurde für Studien mit fehlendem Peer-Review für quantitative (Higgins et al., 2019) und qualitative Studien (Spencer, Ritchie, Lewis und Dillon, 2003) hinzugefügt. Die Ergebnisse zeigen, dass der Einschluss von Studien ohne Peer-Review-Verfahren den Korpus erfolgreich, überzeugend und umfassend erweitert. Der Beitrag diskutiert aus methodischer Sicht, ob und wenn ja, wie sich die verschiedenen Arten von Reviews im Rahmen der Qualitätssicherung und des Publikationsbias unterscheiden und wie sie für den Bereich der empirischen Bildungsforschung bewertet werden können. Insgesamt lässt sich für den Bildungskontext zusammenfassend feststellen,

dass ein vollständiges Bild des Forschungsstand nur möglich ist, wenn auch Zeitschriftenartikel ohne Peer-Review und graue Literatur sichtbar und berücksichtigt werden.

43. Haben nicht-wissenschaftliche Quellen einen Mehrwert? Plädoyer für eine umfassende Suchstrategie bei der Erstellung von Systematic Reviews in der Bildungsforschung

Jolika Schulte, Karin Zimmer
Universität Vechta, Deutschland

Theoretischer Hintergrund und Fragestellung

„Wissenschaftliche Beratung auf der Grundlage des bestmöglichen verfügbaren wissenschaftlichen Wissens“ (Wissenschaftsrat, 2016, S.17) – diese Forderung nach Evidenzorientierung hat den Transfer von wissenschaftlichen Erkenntnissen aus der Forschung in die Gesellschaft, Kultur, Wirtschaft und Politik stark geprägt. Dabei gelten Forschungssynthesen auf Grund der strengen wissenschaftlichen Kriterien, nach denen sie erstellt werden, als besonders geeignet, um alle relevanten Forschungsergebnisse zu einem Themenfeld zu identifizieren, entsprechend ihrer Aussagekraft zu aggregieren und auf dieser Grundlage aus Einzelbefunden Evidenz abzuleiten (vgl. Witte, Woerfel und Schöneberger, 2022, S.114). Dem Anspruch, möglichst alle relevanten Forschungsergebnisse in die Synthese einzubeziehen, kann nur durch eine umfängliche und wohlstrukturierte Suche genügt werden. Neben den quellenspezifischen Suchfunktionalitäten sind deswegen vor allem auch alle relevanten Quellenarten zu berücksichtigen. Dabei herrscht Uneinigkeit darüber, ob neben den in wissenschaftlichen Fachzeitschriften publizierten Forschungsergebnissen dazu auch andere Quellenarten, wie etwa Werke der sogenannten „Grauen Literatur“ zählen.

Ein wichtiges Ziel sowohl von Systematic Reviews im engeren Sinne als auch für einen schnelleren Überblick erstellte Scoping Reviews und sog. Evidence and Gap Maps ist es, neben der Einordnung und Bündelung des verfügbaren Wissens auch, Forschungslücken aufzudecken (vgl. Petticrew and Roberts, 2006; Gough, Oliver and Thomas, 2017). Dies impliziert eine möglichst umfassende Suche, denn nur so kann sichergestellt werden, dass etwaige als Lücken identifizierte Bereiche auf noch ausstehende Forschung schließen lassen und nicht auf vorhandene, aber im Rahmen der eingeschränkten Suche nicht berücksichtigte Arbeiten.

Hier setzt dieser Beitrag an: Am Beispiel eines Systematic Reviews zum Thema institutionelle Sprachförderung von Kindern über drei Jahren in Kindertageseinrichtungen in Deutschland (Zimmer et al. in Vorbereitung) wird demonstriert, wie wichtig es ist, auch in der Wissenschaft eher unübliche Quellen bei der Erstellung eines Systematic Reviews zu nutzen.

Methode

Als Grundlage dient eine systematische Literatursuche die zur Bearbeitung der Frage nach institutioneller Sprachförderung von Kindern in Kindertageseinrichtungen in Deutschland entwickelt wurde (Zimmer et al., in Vorbereitung). Sie hatte zum Ziel, möglichst alle relevanten Publikationen/Untersuchungen aus dem betrachteten Zeitraum 1949 bis Februar 2017 zu finden. Dazu wurde zum einen in insgesamt neun nationalen und internationalen Fachdatenbanken und Projekt-Datenbanken recherchiert. Zum zweiten wurden Expert*innen zum Thema aus Wissenschaft, Verwaltung und Zivilgesellschaft angeschrieben und direkt um Hinweise auf Programme,

Publikationen und nicht veröffentliche Dokumentation zum Thema gebeten. Zum dritten wurde eine ergänzende Internetrecherche über Google und Google Scholar durchgeführt.

Alle Suchabfragen wurden detailliert dokumentiert, um die Suche für Dritte transparent und reproduzierbar zu halten. Insgesamt erbrachte die Suche 2.566 verschiedene Dokumente, von denen 327 als für die Fragestellung relevant in den Systematic Review aufgenommen wurden. Für jedes gefundene Dokument wurde dabei verzeichnet, in welcher Quelle oder Quellen und wie häufig es insgesamt gefunden wurde. Diese Informationen stellen die Grundlage für die im Beitrag vorgestellten Auswertungen dar.

Ergebnisse und ihre Bedeutung

Die Analyse der Nützlichkeit der Quellenarten zum Auffinden dieser 327 Dokumente zeigt, welche Effekte verschiedene Limitationen der Suche (z.B. auf bestimmte Fachdatenbanken, publizierte Dokumente, nur Dokumente aus wissenschaftlichen Zeitschriften mit Gutachtersystem) auf das Suchergebnis haben. Unterschieden wird zwischen Dokumenten, die von der Existenz einer Sprachförderung berichten (insgesamt 327 verschiedene) und denen, die zusätzlich aussagekräftige Angaben zur Wirksamkeit der Sprachförderung machen (31 Fälle). Im Systematic Review zur Sprachförderung sind besonders letztere, die eine mittlere bis hohe Aussagekraft (Weight of Evidence) aufweisen, von großem Interesse.

So zeigt sich zum Beispiel, dass 38% aller relevanten Dokumente über das direkte Nachfragen bei Expert*innen und Institutionen und die ergänzende Internetrecherche gefunden wurden. Die zur Wirksamkeit von Sprachförderung herangezogenen Untersuchungen stammen knapp zur Hälfte aus diesen Quellen, 8 von 31 wurden sogar ausschließlich dort gefunden. Die Ergebnisse werden im Hinblick auf die Konzeption von Suchstrategien für Forschungssynthesen weiter diskutiert.

44. Die Untersuchung von Gruppenverzerrungen bei der automatischen Kodierung offener Textantworten in Large-Scale Assessments

Nico Andersen[1], Fabian Zehner[1,2], Julia Mang[2,3], Frank Goldhammer[1,2]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; [2]Zentrum für Internationale Bildungsvergleichsstudien (ZIB); [3]Technische Universität München

Die hier präsentierte Forschungsarbeit setzt sich damit auseinander, ob automatische Kodiersysteme Entscheidungen bezüglich verschiedener Gruppen verzerren und wenn ja, unter welchen Bedingungen, wobei diese auf unterschiedlichen Ebenen analysiert werden.

Durch den wachsenden Fortschritt wird die Automatisierung menschlicher Arbeit immer mehr in den Alltag integriert. Automatische Systeme werden dabei auch im Bildungsbereich, unabhängig von ihrer Komplexität, häufig an wichtigen Entscheidungen beteiligt. Eine besondere Herausforderung ist die Kodierung von offenen Textantworten, welche aufgrund der sprachlichen Komplexität aktuell häufig noch manuell ausgeführt wird. Mittels maschinellen Lernen und natürlicher Sprachverarbeitung werden jedoch stetig bessere automatische Kodierergebnisse erzielt (s. Yan, Rupp & Foltz, 2020), was einen breiten Einsatz attraktiver macht.

Bei der Methodenentwicklung steht die durchschnittliche Kodiergenauigkeit als Evaluationsmaß im Mittelpunkt, wobei Verzerrungen (systematische Fehlkodierungem zum Vor- oder Nachteil verschiedener Gruppen) häufig nicht erfasst werden.

Die Methoden unterscheiden sich in ihren Ansätzen. Es gibt beispielsweise Clustermodelle, welche besonders transparent hinsichtlich ihrer Kodierentscheidungen sind (s. Zehner, Sälzer & Goldhammer, 2016), oder Modelle, welche auf Basis eines tiefen neuronalen Netzes funktionieren und als „Black Box" bezeichnet werden (s. Devlin, Chang, Lee & Toutanova, 2019). Bei modernen Methoden werden Antworten in einem semantischen Raum abgebildet, in dem Antworten auf Basis der Kosinus-Ähnlichkeit ihrer Vektoren als semantisch ähnlich identifiziert werden können. Mittels dieses Maßes und einem Trainingsdatensatz können weitere Antworten automatisch klassifiziert werden.

Aus der Vielfalt anwendbarer Methoden ergibt sich auch die Fragestellung, inwiefern sich Verzerrungen hinsichtlich der gewählten Methode unterscheiden. Untersuchungen zu automatischen Kodierungen im Kontext des Spracherwerbs konnten beispielsweise Verzerrungen zu Gunsten bestimmter Gruppen nachweisen (Bridgeman, Trapani & Attali, 2009).

Als Datenbasis unserer Analysen dienen die Antworten der Lesefähigkeitsmessung des *Programme for International Student Assessment* (PISA) aus dem Erhebungsjahr 2018 (Weis et al., 2018). Diese lassen Gruppenvergleiche innerhalb einer Sprache, zwischen Personen unterschiedlichen Geschlechts sowie einen internationalen Vergleich zwischen den Testsprachen zu. Es werden insgesamt 30 Items mit durchschnittlich 919 Antworten pro Item und einer durchschnittlichen Antwortlänge von M=15.8 Tokens (Wörter und Zahlen; SD=16.5) untersucht, wobei ausschließlich Items mit dichotomer Bewertung (richtig/falsch) berücksichtigt werden.

Das Untersuchungsmaß für Verzerrungen wird durch die Kodiergenauigkeit abgebildet (die Anzahl der relativen Übereinstimmungen zwischen automatischen und manuellen Kodierungen). Die Differenz der durchschnittlichen Kodiergenauigkeit zweier Gruppen bildet die Verzerrung ab.

Aktueller Stand der Untersuchung:

Erste Untersuchungen wurden für den Vergleich zwischen Jungen und Mädchen durchgeführt. Mittels nicht-parametrischer Tests wurde untersucht, ob Verzerrungen von der Gruppeneinteilung abhängig sind. Dafür wurden Kodierungen mittels Bootstrapping zufällig einer Gruppe zugeordnet. Es konnten dabei keine generischen systematischen Unterschiede gefunden werden, jedoch systematische Unterschiede auf der Item-Ebene zugunsten bzw. zulasten einzelner Gruppen.

Ausblick:

Das Ziel der Studie, welche bis zum Zeitpunkt der Präsentation abgeschlossen sein wird, ist es, weitere Kodiermethoden zu testen (z.B. Deep Learning) sowie Gruppenvergleiche zwischen Muttersprachler*innen und Nicht-Muttersprachler*innen durchzuführen und diese Ergebnisse auf Methoden- sowie Gruppen- und Itemlevel gegenüberzustellen. Des Weiteren sollen Gründe für Verzerrungen exploriert werden, wie beispielsweise gruppenspezifische linguistische Merkmale, welche unterschiedliche Ergebnisse bei der Anwendung verschiedener Methoden bedingen können.

45. Textanreicherungsaufgaben als neues Aufgabenformat im NEPS-Lesekompetenztest

Kathrin Thums, Karin Gehrer
LIfBi – Leibniz-Institut für Bildungsverläufe, Deutschland

Hintergrund:

Neben vielfältigsten Erhebungsinhalten untersucht das Nationale Bildungspanel (NEPS) unter anderem die Entwicklung verschiedener Kompetenzen über die gesamte Lebensspanne (Weinert et al., 2019). Eine essentielle Kompetenz, welche im NEPS über alle Startkohorten hinweg im Längsschnitt und kohärent untersucht wird, ist die Lesekompetenz der Teilnehmenden. Lesekompetenz meint dabei die funktionale Fähigkeit geschriebene Texte zu verstehen (Gehrer et al., 2013). In Hinblick auf die Erhebungen der Lesekompetenz ermöglicht eine Umstellung von papierbasierten auf computergestützte Tests (CBA) die Implementierung neuer Aufgabenformate. Bislang wurden im NEPS-Lesekompetenztest drei Arten von Aufgabenformaten zur Messung der Lesekompetenz eingesetzt: einfache Multiple-Choice-Aufgaben (MC; eine richtige Antwort aus vier Antwortoptionen), Entscheidungstabellen (CMC; richtig/falsche Antworten oder stimmt/stimmt nicht Optionen) und Zuordnungsaufgaben (ZO; Gehrer et al., 2012). Als innovatives Aufgabenformat werden in CBA-Lesekompetenzmessungen Lückentext-Aufgaben bzw. Textanreicherungsaufgaben (TAA) eingesetzt (Rohm et al., 2019). Das neue TAA-Format innerhalb des NEPS-Lesekompetenztests wurde entwickelt in Anlehnung, aber auch in Abgrenzung zum klassischen Format von Lückentexten in Cloze- oder C-Tests (z.B. Eckes 2010). Da bei Cloze-Tests Wörter und den C-Tests Worthälften systematisch entfernt werden, stellen diese Lückentexte „zerstörte Nachrichten“ (Klein-Braley, 1997) für die Lesenden dar. Im Unterschied dazu handelt es sich beim TAA-Aufgabenformat des NEPS-Tests um vollständige Texte. Auf der Basis des gewonnenen Situationsmodelles (Kintsch, 1998) ist das Ziel des neuen Formates eine sinnvolle Anreicherung des Textes zu erreichen, durch eine zusätzliche Platzierung von ganzen Sätzen. Im Unterschied zum MC-Cloze-Test (Jonz, 1976) werden jedoch nicht mehrere Optionen für eine Lücke angeboten, sondern es werden beim TAA-Format mehrere Lücken für eine Option angeboten. Die Aufgabe der Lesenden bei der TAA besteht darin, den vorgelegten Text mit drei bis vier zusätzlichen Sätzen sinnvoll zu ergänzen. Dies bedeutet, dass sie die richtigen Stellen im Text finden müssen, an denen die zusätzlichen Sätze am besten passen, sowohl semantisch als auch grammatikalisch. Die Herausforderung für die erfolgreiche Bearbeitung der TAA-Aufgaben besteht darin, die Aussagen des Textes, die Argumentation oder die im Text beschriebene Handlung zu verstehen als auch die möglichen Verknüpfungen mit den Kontexttextstellen zu prüfen, bevor eine Entscheidung über die korrekte Position des zusätzlichen Arguments getroffen werden kann. Da zur Lösung der TAA-Aufgaben zunächst ein Situationsmodell des Textes und seiner Abschnitte als auch eine entsprechende Repräsentation mehrerer möglicher Textstellen erarbeitet werden muss, stellen sie insbesondere für schwächere Lesende eine besondere Herausforderung dar. In ihrer Komplexität sind TAA innerhalb eines Testlets eine valide Ergänzung bisheriger MC-, ZO- und CMC-Formate zu beliebigen Textsorten.

Fragestellung:

Um die Qualität der erstmalig eingesetzten TAA-Aufgaben in einer NEPS-Hauptstudie zu untersuchen, werden im vorliegenden Beitrag Berechnungen zur Passung der Aufgaben durchgeführt. Folgende Fragen werden untersucht: Wie schneiden verschiedene Lesende bei TAA-Aufgaben ab und lassen sich Unterschiede zwischen guten und schlechten Lesenden identifizieren?

Methode:

Die TAA-Aufgaben in einem NEPS-Lesekompetenztest wurden erstmals in einer NEPS-Hauptstudie mit insgesamt 17.851 Teilnehmenden (53 % Frauen) aus drei verschiedenen Startkohorten (SC4 *N* = 6.863; SC5 *N* = 4.591; SC6 *N* = 6.397) in einer verlinkten Erhebung eingesetzt. In den fünf bearbeiteten Texten aus verschiedenen Textsorten (Sachtext, literarischer Text, Anleitung, Kommentar, Aufruf/Werbung) haben die Teilnehmenden TAA-Aufgaben bearbeitet.

Ergebnisse und Bedeutsamkeit:

Erste Ergebnisse aus multivariaten Gruppenvergleichen zeigen, dass gute Lesende und jüngere Studienteilnehmende häufiger die TAA-Aufgaben richtig lösen als dies bei älteren und schlechter Lesenden der Fall ist. Aufgrund der insgesamt heterogenen Altersspanne der Teilnehmenden (19-73 Jahre) werden weitere Analysen nochmal getrennt für Altersgruppen und Bildungshintergrund durchgeführt. Da es sich um eine computer-basierte Erhebung handelt, könnte neben der Lesekompetenz auch der Umgang mit der Technik eine Rolle spielen, insbesondere bei älteren Teilnehmenden. Daher werden die Ergebnisse nicht nur im Hinblick auf die Lesekompetenz der Teilnehmenden diskutiert, sondern es wird auch die Messung der Lesekompetenz näher betrachtet.

46. Effects of mergers on the financing and scope of education offered by Adult Education Centres in Germany

Nicolas Echarti, Elisabeth Reichart, Andreas Martin
Deutsches Institut für Erwachsenenbildung Leibniz-Zentrum für Lebenslanges Lernen e. V., Germany

Reforms to consolidate public education providers through institutional mergers are part of government reform policies in many countries (see e.g., Azziz et al., 2019; De Boer et al., 2017; Pinheiro et al., 2017). However, the impact of such reforms on the scope, costs and cost-efficiency of educational services provided is often uncertain. On the one hand, mergers can lead to efficiencies that allow institutions to provide more and better-quality services (see e.g., Russel, 2019; Liu et al., 2018; Harman & Harman, 2003), on the other hand, mergers can also negatively impact efficiency and institutional output (cf. Slade et al., 2022; Johnes & Tsionas, 2018; Beuchert et al., 2018; Stensaker et al., 2016).

Our study provides an empirical example of estimating merger effects from historical data using Poisson regression and robust standard errors. As shown by the considerable number of documented merger cases (Reichart, 2018, p. 189), institutional amalgamation also plays a key role in the field of publicly funded continuing education in Germany. Research to date on the effects of mergers in continuing education in Germany has been limited to specific contexts and merger scenarios (Mickler, 2013; Weißenberg & Sedlack, 2007). We add to the literature by using extensive longitudinal data to analyse the extent to which mergers between adult education centres in Germany (Volkshochschulen, VHS) have affected funding, educational offerings, and the financial efficiency of service delivery post-merger.

To analyse the impact of institutional mergers on the provision of continuing education, we supplemented historical revenue and output data (up to 24 waves, from 1994 to 2017) from 870 VHS (DIE, 2018) with detailed information on historical merger events (n=92). We then used Poisson pseudo-maximum likelihood regression with multiple levels of fixed effects and cluster-robust variance estimation (Correia et al., 2020) to examine the relationship between the mergers and outcome

variables over time. The fixed effects comprised institutional fixed effects, year fixed effects and panel-specific slopes. The outcomes examined were annual funding, annual teaching hours, and funding per ten-thousand hours taught. Funding was further differentiated by the funding source, i.e. participation fees, local funds, state funds as well as federal and EU funds. Using a multivariate indicator variable to capture temporal variations between pre- and post-merger years, we estimated marginal effects of the mergers for the first four years post-merger and for the long-term average (five years or more post-merger); the pre-merger years were used as reference category.

We find that in post-merger periods, total funding was continuously lower compared to the situation before institutional amalgamation, but only the estimate for the first year was statistically significant. Local funding from municipalities and counties was steadily reduced after the mergers, while state, federal and EU funding increased. Participation fees collected did not change considerably. Educational output, measured by the number of annual teaching hours, was temporarily lowered, but none of the estimated coefficients were statistically significant and the long-run estimate of the merger effect was close to zero. Financial efficiency, as measured by the ratio of financial resources used to teaching output, showed temporary fluctuations in the short term, but ultimately also did not change significantly.

In summary, the merger cases studied were characterised by lasting changes in funding channels and temporary disruptions in the financial capacity, teaching output and efficiency - in the long run overall inputs and outputs were not significantly affected by the mergers, and financial efficiency also remained unchanged. Further research on the effects of mergers in continuing education could investigate the impact on other policy-relevant outcomes such as the diversity and quality of course offerings or their reach and accessibility.

47. Diversitätswissen von Lehramtsstudierenden im Kontext einer inklusionsorientierten Lehrkräftebildung - Ergebnisse einer Fragebogenerhebung

Sarah Lamb[1], Sarah Schulze[1], Sven Anderson[1], Jan-Simon Zimmermann[1], Anne Schröter[2], Jörg-Tobias Kuhn[1], Ricarda Steinmayr[1]
[1]Technische Universität Dortmund; [2]Leibniz Universität Hannover

Theoretischer Hintergrund. Die Qualifizierung von Lehramtsstudierenden für die Gestaltung inklusiver Schul- und Unterrichtsprozesse bedarf inklusionsbezogener Inhalte, wie z.B. Kenntnisse über Diversität und Heterogenität als Bedingungen von Schule und Unterricht (KMK, 2014), die lehramtsübergreifend an Universitäten curricular verankert werden müssen (Hußmann, Schlebrowski, Schmidt & Welzel, 2018, KMK, 2014). Allgemein wird in kompetenztheoretischen Professionalisierungsansätzen angenommen, dass die Qualität des Unterrichts (und damit die Kompetenzentwicklung der Schüler*innen) mit der Ausprägung von Kompetenzaspekten (angehender) Lehrkräfte, wie u.a. Wissen, Selbstwirksamkeit und Einstellungen, zusammenhängt (z.B. Baumert & Kunter, 2011). Bislang hat sich die Kompetenzforschung im Kontext der schulischen Inklusion stark auf die Einstellungen und Selbstwirksamkeit fokussiert. Das pädagogisch-psychologische Wissen über heterogene Voraussetzungen der Lernenden (Voss, Kunina-Habenicht, Hoehne & Kunter, 2015) wurde hingegen deutlich seltener beforscht, obwohl angenommen wird, dass der Einbezug von Wissensinhalten eine Basis für die effektive und individuelle Förderung aller

Schüler*innen bildet (z.B. Pohl, 2020). Hinzukommt das ein solches Wissen u.a. in wechselseitiger Beziehung zu inklusionsbezogenen Selbstwirksamkeit, Einstellungen, Lerngelegenheiten zu Diversität steht (Tometten, Heyder & Steinmayr, im Druck).

Da bislang kaum geeignete Instrumente zur Erfassung des Diversitätswissens im Kontext schulischer Inklusion vorlagen, wurde im Rahmen des Projektes „Dortmunder Profil für inklusionsorientierte Lehrer/-innenbildung“ (DoProfiL) ein Wissenstest zu fünf verschiedenen Diversitätsbereichen: Förderschwerpunkt emotional-soziale Entwicklung und Lernen, Geschlecht, kognitive Hochbegabung und Teilleistungsstörungen, entwickelt (Heyder, Vaskova, Hußmann & Steinmayr, 2018). Es gibt Hinweise darauf, dass z.B. Lehramtsstudierende des sonderpädagogischen Lehramtes mehr Fragen des Wissenstests richtig lösen als Studierende allgemeinbildender Schulformen (Tometten et al., im Druck). Ziel der hier vorliegenden Studie war es daher systematisch zu untersuchen, ob das Diversitätswissen im Laufe des Studiums von den Studierenden unterschiedlicher Studienabschnitte und Lehrämter erworben und abgerufen werden kann.

Fragestellung. Untersucht wurde, (1) ob sich die Studierenden verschiedener Lehrämter und Studienabschnitte in ihrem Wissen zu den fünf verschiedenen Diversitätsbereichen unterscheiden und (2) ob, es Unterschiede in den Wissensprofilen innerhalb der verschiedenen Lehrämter und Studienabschnitte gibt.

Methode. Dazu wurde der im Kontext von DoProfiL konzipierte Wissenstest (Steinmayr, Heyder & Tometten, 2022) im Rahmen einer querschnittlichen Fragebogenerhebung mit N = 992 Lehramtsstudierenden (Lehramt an Grundschulen (G) = 23 %, Lehramt an Haupt-, Real-, Sekundar- und Gesamtschulen (HRSGe) = 10 %, Lehramt an Gymnasien und Gesamtschulen (GyGe) = 25 %, Lehramt an Berufskollegs (BK) = 9 % und sonderpädagogische Förderung (SP) = 33 %) eingesetzt.

Ergebnisse. Die Studie liefert relevante Erkenntnisse für die Ausgestaltung inklusionsorientierter Lehr- und Lernangebote. Die wichtigsten Ergebnisse waren, dass sich die SP-Studiereden curricular erwartbar von den Studierenden allgemeinbildender Schulformen abhoben, indem sie zu allen Diversitätsbereichen höhere Wissensausprägungen hatten. Theoriekonform (vgl. Qualifikationshypothese, z.B. Darling-Hammond, 2006) war das Wissen der Masterstudierenden höher ausgeprägt als das Wissen der Bachelorstudierenden. Die Analyse der Wissensprofile zeigte, dass das Wissen innerhalb der Bachelor- und Masterkohorte zu den fünf verschiedenen Diversitätsaspekten nicht substanziell unterschiedlich stark ausgeprägt war. Innerhalb einzelner Lehrämter zeichneten sich jedoch Unterschiede im Diversitätswissen ab. Beispielsweise dominierte lehramtsbedingt erwartbar innerhalb der SP-Gruppe das Wissen zu den Förderschwerpunkten emotional-soziale Entwicklung und Lernen, wohingegen in den Gruppen allgemeinbildender Schulformen das Wissen zum Förderschwerpunkt Lernen vergleichsweise niedrig war. Die Studie identifiziert Diversitätsbereiche die im Kontext der Professionalisierung stärker berücksichtigt werden sollten, um die Lehramtsstudierenden in ihrer Wissensentwicklung zu unterstützen und damit auf die inklusionsorientierte Unterrichtsgestaltung vorzubereiten. Beispielsweise lösten die G-Studierenden trotz hoher Prävalenzen (z.B. Fischbach et al., 2010) im Mittel nur 34 % der Fragen zum Diversitätsmerkmal Teilleistungsstörungen korrekt. Die Ergebnisse liefern damit wichtige Hinweise für

die inhaltliche Ausgestaltung künftiger Lehr- und Lernangebote. Im Hinblick auf das Grundschulstudium könnte z.B. Handlungsbedarf diskutiert werden, da Teilleistungsstörungen meist im Grundschulalter mit den Leistungsbewertungen in der Grundschule erstmalig in Erscheinung treten (WHO, 2021) und daher die Aneignung von fundiertem Wissen zu Teilleistungsstörungen bereits während des Studiums bedeutsam wäre.

48. Lassen sich Ausprägungen professionsunabhängiger Einstellungen zu einem inklusiven Schulsystem in einer Stichprobe von Studierenden der Geographiedidaktik replizieren?

Dietmar Gölitz, Ann-Sophie Winklmaier, Christiane Tellesch-Bülow, Jan Christoph Schubert
Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Empirischer Hintergrund

Im Rahmen eines evidenzbasierten Forschungsprojekts zur Qualifizierung von Geographielehrpersonen für Inklusive Bildung (Winklmaier et al., 2022) erhoben wir mit Studierenden der Geographiedidaktik fragebogenbasiert unter anderem allgemeine Einstellungen zum inklusiven Schulsystem (PREIS; Lüke & Grosche, 2017), denn aus geographiedidaktischer Perspektive ist interessant, mit welchen professionsunabhängigen Eingangsvoraussetzungen Studierende das Themenfeld Inklusion aus der geographiedidaktischen Perspektive bearbeiten. Spezifischen positiven Einstellungen wird als Teil der professionellen Handlungskompetenzen von (angehenden) Lehrpersonen (Baumert & Kunter, 2006) im Kontext von Inklusion eine besondere Bedeutung zugeschrieben (Loreman, Forlin & Sharma, 2014) und nationale und internationale Forschungsarbeiten (Heyl & Seifried, 2014; Lüke & Grosche, 2018; Seifried, 2015; Seifried & Heyl, 2016; Sharma & Nuttal, 2016) haben sich mit Einstellungen zu inklusivem Unterricht auseinandergesetzt. Um einen Bezug zu spezifischeren Maßen pädagogisch relevanter Einstellungen herzustellen, wäre es wünschenswert, ein allgemeines reliables Verfahren einzusetzen. Daher soll hier die Eignung des Verfahrens PREIS geprüft werden. Dazu werden Daten zu diesen Einstellungen berichtet und mit Studien der Skalenautor:innen verglichen (Lüke & Grosche, 2018; Schulze, Lüke, Schröter, Krause & Kuhl, 2019). Ziel der Analyse ist, im Rahmen einer konzeptuellen Replikation zu prüfen, mit welcher Anzahl von Faktoren die interindividuellen Unterschiede in den Antworten der befragten Stichprobe inhaltlich am besten beschrieben werden können. Der Rohdatensatz für die Stichprobe wurde im Rahmen der Projektberichterstattung veröffentlicht (Gölitz et al., 2021).

Fragestellung

Inwieweit lassen sich professionsunabhängige Einstellungen zu einem inklusiven Schulsystem in einer Stichprobe von Studierenden der Geographiedidaktik als eindimensional beschreiben?

Methode

Wir prüfen über die die interne Konsistenz und Faktorenstruktur (IBM, ohne Jahr), inwieweit es sich um ein eindimensionales Konstrukt handelt (N = 276 ohne fehlende Daten). Eine konzeptuelle Veränderung besteht in der der verwendeten Rating-Skala, die von einer vierstufigen auf eine sechsstufige Skala erweitert wurde.

Resultat

Es resultieren klar positive Zustimmungsraten zu einem inklusiven Schulsystem. Diese liegen mit teilweise großen Effekten über dem rechnerischen Skalenmittelwert von 3.5. Im Vergleich zu aus der PREIS-Forschung berichteten Daten resultieren minimal voneinander abweichende Mittelwerte für die einzelnen Items. Das Profil der Mittelwerte für die 14 Items entspricht damit den Vergleichsstichproben. Cronbachs Alpha beträgt .92. Die dazugehörige Faktorenanalyse mit der Vorgabe von einem Faktor weist allerdings nur einen unbefriedigenden Fit auf (Chi-Quadrat = 255, df = 77, p < .00). Mit einer Faktorenanalyse mit Eigenwerten > 1 ergeben sich zwei Faktoren (Chi-Quadrat = 169, df = 64, p < .00), ein Generalfaktor mit 51% Varianzaufklärung und ein Faktor für die vier gegensinnig zur positiven Einstellung formulierten Items des PREIS. Da der Scree-Plot auf fünf Faktoren hinweist, berechneten wir eine Analyse mit fünf Faktoren. Die fünf-faktorielle Lösung ergibt den besten Fit dieser Faktorenanalysen (Chi-Quadrat = 47, df = 31, p = .03) und trennt zusätzlich als interpretierbare Facetten Handlungsabsichten (bspw. eine Partei gegen ein inklusives Schulsystem wäre nicht wählbar) und ein Item zu Schulunterricht von den beiden Facetten der zweifaktoriellen Lösung ab.

Diskussion

Obwohl hier im Gegensatz zu den Bezugsstudien eher Hinweise auf die Mehrdimensionalität des PREIS vorliegen, scheint es gerechtfertigt, wie die Originalautor:innen (Schulze et al., 2019), einen eindimensionalen Gesamtscore zu verwenden, da sich eine sehr gute Reliabilität ergibt. In diesem Zusammenhang kann nach den hier vorgestellten Analysen eine Kurzskala mit ca. der Hälfte der Items verwendet werden. Aus Sicht des Ansatzes von Replikationen (Gelman, A., Hill & Vehtari, 2021; Gelman, A. & Loken, 2014) macht es Sinn, vorhandene Verfahren kontinuierlich anhand neuer Daten zu reanalysieren. In der Gesamtbewertung sind die über die PREIS-Skala erhobenen Einstellungen zu einem inklusiven Schulsystem in der erhobenen Stichprobe als Gesamtscore für interindividuelle Differenzen zufriedenstellend abgebildet. Die Replikation einer eindimensionalen Faktorenstruktur scheint gelungen. Wir diskutieren die Potenziale und Grenzen der Verwendung erprobter Verfahren für die geographiedidaktische Evaluations- und Transferforschung aus Sicht der internalen Validität.

49. Stereotype von Lehrkräften über Schülerinnen und Schülern mit Migrationshintergrund

Helene Zeeb[1], Thamar Voss[2]
[1]Universität Erfurt, Germany; [2]Albert-Ludwigs-Universität Freiburg

Das Denken und Handeln von Lehrkräften kann durch Stereotype – d.h., verallgemeinerte Annahmen über eine soziale Gruppe – verzerrt sein. Beispielsweise können negative Stereotype über Lernende mit Migrationshintergrund mit ungünstigen Einschätzungen einhergehen (z.B. Froehlich et al., 2015; Glock & Böhmer, 2018). Die Erfassung von Stereotypen ist jedoch in mehrfacher Hinsicht herausfordernd: Erstens werden die Inhalte der Stereotype meist vorab festgelegt, zum Beispiel durch Fragebogen-Items oder IAT-Wortlisten (z.B. Glock & Böhmer, 2018; Hachfeld et al., 2012) und sind daher durch Vorannahmen der Forschenden bestimmt. Zweitens werden Stereotyp-Inhalte, wenn sie frei erfasst werden, oftmals in allgemein-gesellschaftliche Bereiche unterteilt (z.B. Froehlich et al., 2015), welche jedoch für den Schulkontext wenig aussagekräftig sind. Drittens werden die Personen üblicherweise nicht nach ihrer persönlichen Meinung gefragt, sondern danach, welche Annahmen

ihrer Ansicht nach in der Gesellschaft vorherrschen (z.B. Bonefeld & Karst, 2020; Froehlich et al., 2015). Dieses Vorgehen birgt die Gefahr, dass die erfassten Stereotype viel stärker sind als die tatsächlich vorhandenen (Kotzur et al., 2020). Ziel der vorliegenden Studie war es daher, die (subjektiv vorhandenen) Stereotype von Lehrkräften über Lernende mit Migrationshintergrund offen zu erfassen und die Inhalte nach schulrelevanten Bereichen zu systematisieren. Dieser Systematisierung legten wir das Erwartungs-Wert-Modell der Lernmotivation (Eccles & Wigfield, 2002) zugrunde. Die übergreifende Forschungsfrage lautete: In welche schulrelevanten Inhaltsbereiche lassen sich die Stereotype von Lehrkräften über Lernende mit Migrationshintergrund untergliedern?

An der Befragung nahmen 186 praktizierende Lehrkräfte teil (Alter: *M* = 43.4 Jahre, *SD* = 10.3; 72% weiblich), welche im Durchschnitt 14.9 Jahre (*SD* = 9.6) im Beruf tätig waren. Die Lehrkräfte arbeiteten vor allem an Gymnasien (44%), Real- und Hauptschulen (41%), Gemeinschaftsschulen (13%) und Grundschulen (13%). Der durchschnittliche Anteil an Lernenden mit Migrationshintergrund in den unterrichteten Klassen lag bei 26% (*SD* = 23%). Zur Erfassung der Stereotype wurden zwei freie Produktionsaufgaben verwendet: „Welche (a) positiven Aspekte und (b) negativen Aspekte sehen Sie beim Unterrichten von Schülerinnen und Schülern mit Migrationshintergrund, vor allem hinsichtlich Motivation und Lernverhalten?“ Die Lehrkräfte notierten ihre Antworten in ein leeres Textfeld; die Befragung war online und anonym. Die Auswertung erfolgte mittels qualitativer Inhaltsanalyse. Im ersten Schritt wurden die Antworten nach Inhaltsbereichen und Valenz (positiv/negativ) sortiert. Diese Inhaltsbereiche wurden von einer zweiten Person geprüft und dem Modell der Lernmotivation zugeordnet.

Die wichtigsten Inhaltsbereiche waren:

- Kultureller Hintergrund, z.B. traditionelle Ansichten über die Rolle der Frau (negativ)
- Einstellungen und Verhaltensweisen bedeutsamer Bezugspersonen, z.B. hoher Stellenwert der Bildung bei den Eltern (positiv)
- Vorangegangene Ereignisse, z.B. Fluchterfahrung (negativ)
- Selbstbild und persönliche Ziele, z.B. Ziel der gesellschaftlichen Teilhabe (positiv)
- Emotionen, z.B. Frustration und Langeweile (negativ)
- Subjektive Wertüberzeugungen, z.B. Bildung ist wichtig für die Zukunft (positiv)
- Subjektive Erfolgserwartung, z.B. geringe Einschätzung der eigenen Fähigkeiten, „Ich kann das eh nicht“ (negativ)
- Lernverhalten, z.B. Disziplin und Anstrengungsbereitschaft (positiv)

Zudem wurden Häufigkeiten analysiert, um zentrale Stereotype zu identifizieren. Positive Aspekte wurden am häufigsten in der Kategorie *Lernverhalten* genannt (*n* = 101), und zwar dahingehend, dass (insbesondere weibliche) Lernende mit Migrationshintergrund mit viel Fleiß und Ehrgeiz lernen. Negative Aspekte wurden am häufigsten in der Kategorie *Einstellungen und Verhaltensweisen bedeutsamer Bezugspersonen* genannt (*n* = 87), und zwar vor allem dahingehend, dass die Eltern ihre Kinder wenig beim Lernen unterstützen.

Die Ergebnisse bieten einen Überblick über lernbezogene Stereotype von Lehrkräften über Lernende mit Migrationshintergrund. Eine besondere Stärke der Studie ist, dass Stereotype unbeeinflusst von vorgegebenen Kategorien erfasst wurden. Einschränkungen ergeben sich dadurch, dass die explizite Erfragung möglicherweise sozial erwünschte Antworten produzierte und dass bei der Zielgruppe

(Lernende mit Migrationshintergrund) nicht genauer nach Herkunftsregionen differenziert wurde. Die Ergebnisse können als Grundlage für die Entwicklung von Messinstrumenten dienen, welche speziell auf den Schulkontext zugeschnitten sind. Darüber hinaus sollten Ideen entwickelt werden, wie verbreitete, nichtzutreffende Stereotype entkräftet werden könnten.

50. Wissen und Einstellungen von Lehrkräften der Sekundarstufe I zu Mehrsprachigkeit und Familiensprachen in der Schule

Natalia Robak, Robert Reggentin, Katrin Böhme
Universität Potsdam, Deutschland

Das Thema Mehrsprachigkeit in der Schule ist höchst relevant. Zuletzt erhielt die Frage danach, ob die schulische Entwicklung geflüchteter Kinder und Jugendlicher aus der Ukraine in separaten „Willkommensklassen" oder in Regelklassen besser gelingt und welchen Stellenwert die Vermittlung der deutschen Sprache für diese Lernenden hat, hohe mediale Aufmerksamkeit. Während im deutschen Bildungssystem Einsprachigkeit lange als Normalfall angesehen wurde, wird heute Mehrsprachigkeit als selbstverständliche Alltagsrealität wahrgenommen (Busch, 2017). Dennoch dominiert in Deutschland das Modell des einsprachigen Schulunterrichts (Busch, 2017) und mehrsprachige Schüler:innen werden überwiegend in der Unterrichtssprache Deutsch gefördert. Für Kinder mit Deutsch als Zweitsprache bedeutet das, dass sie ihre Familiensprachen nur sehr begrenzt in den Unterricht einbringen können, obwohl wissenschaftliche Erkenntnisse die Argumentation für den Eingang der Familiensprachen in die Schule und dessen Förderung im schulischen Kontext maßgeblich bestärken (bspw. Gantefort, 2013; Roll et al., 2019; Usunova & Schnoor, 2021).

Sowohl für die Gestaltung von Unterrichtsprozessen als auch für den gelingenden Umgang mit Mehrsprachigkeit in sprachlich heterogenen Klassen wird in der Fach- und Forschungsliteratur auf die Bedeutung vom fachdidaktischen Wissen sowie von Überzeugungen und Einstellungen von Lehrkräften hingewiesen (Hachfeld, 2015; Hirschauer & Kullmann, 2010; Michalak, 2010; Trautmann & Wischer, 2011). Nach Baumert und Kunter (2006) entsteht professionelle Handlungskompetenz aus dem Zusammenwirken von deklarativem und prozeduralem Wissen, professionellen Überzeugungen und Werthaltungen, motivationalen Orientierungen sowie Fähigkeiten zur Selbstregulation. Da das Wissen und die Überzeugungen bzw. die Einstellungen als eng aufeinander bezogene Konstrukte gelten, stellt sich die Frage, ob das Wissen der Lehrkräfte zu wissenschaftlich fundierten Theorien zur Mehrsprachigkeit in Beziehung zu ihrer Überzeugung zur Förderung der Familiensprachen und ihrer Einstellung zur Nutzung von Angeboten in Familiensprachen in der Schule steht und welche Bedeutung relevante Hintergrundvariablen (Berufserfahrung, Ausbildung sowie eigene Spracherfahrungen) für die genannten Konstrukte (Wissen und Einstellungen) haben. Da die wissenschaftlichen Studien zum Einfluss des Wissens auf die Einstellungen teilweise zu widersprüchlichen empirischen Befunden kommen (Fröhlich-Gildhoff et al., 2014; Kratzmann et al., 2017, 2020), wurde ein Online-Fragebogen für Lehrkräfte der Sekundarstufe zu mehrsprachigkeitsbezogenem Wissen und Einstellungen zu Familiensprachen in der Schule entwickelt.

An der Fragebogenstudie haben N = 86 Lehrkräfte (w:72%, m:25%, d:2%) aus Berlin und Brandenburg teilgenommen. Die Stichprobe ist bezüglich des Alters annähernd normalverteilt (18-24

Jahre = 1,2%, 25-34 Jahre = 29,3%, 35-44 Jahre = 31,7 %, 45-54 Jahre = 18,3%, über 54 Jahre = 19,5%). Im Mittel verfügen die beteiligten Lehrkräfte über mehr als ein Jahrzehnt Berufserfahrung (M = 12,7 Jahre, SD = 11,0 Jahre). Alle Skalen zeigen eine mindestens befriedigende Reliabilität α ≥ 0,70. Die befragten Lehrkräfte sind gegenüber der Förderung von Familiensprachen in der Schule positiv eingestellt und vertreten mehrheitlich die Meinung, dass schulische Lehr- und Freizeitaktivitäten, wenn möglich auch in Familiensprachen angeboten werden sollten. Zwischen dem mehrsprachigkeitsbezogenen Wissen und der Einstellung zur Förderung der Familiensprachen sowie zu Angeboten in den Familiensprachen zeigen sich positive Zusammenhänge in Höhe von r = 0,38-0,64. Es zeigte sich auch, dass Lehrkräfte, deren Erstsprache nicht Deutsch ist, über ein größeres mehrsprachigkeitsbezogenes Wissen sowie eine positivere Einstellung zu Familiensprachen in der Schule verfügen. Zwischen Wissen und Einstellungen einerseits und den Hintergrundvariablen Berufserfahrung sowie Ausbildung andererseits konnte kein Zusammenhang identifiziert werden.

Die Ergebnisse stehen im Einklang mit den in der Fachliteratur verankerten Professionalisierungstheorien von Lehrkräften im Umgang mit mehrsprachigen Kindern im schulischen Kontext und liefern einen empirischen Beitrag zur Mehrsprachigkeitsforschung. Ferner liefern die Ergebnisse Praxisimplikationen für die Professionalisierung der Lehrkräfte: Besonders für Lehrkräfte mit der Erstsprache Deutsch sollten Aus- und Weiterbildungen angeboten werden, die verdeutlichen, welchen Mehrwert eine stärkere Einbeziehung von Familiensprachen der Schüler:innen in den schulischen Alltag haben und wie ein erfolgreiches Lernen in mehrsprachigen Klassen gelingen kann.

51. Lehrkraftreaktionen auf soziale Ausgrenzung unter Schüler*innen

Katharina Fink[1], Jeanine Grütter[2], Hanna Beißert[1,3]
[1]DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; [2]Universität Konstanz; [3]Goethe Universität Frankfurt

Theoretischer Hintergrund

Wenn Kinder regelmäßig soziale Ausgrenzung erleben, kann dies schwerwiegende Konsequenzen haben (e.g., Gazelle & Druhen, 2009; Sebastian et al., 2010). Es ist wichtig zu untersuchen, wie Lehrkräfte auf soziale Ausgrenzung reagieren, da sie mit ihren Reaktionen Vorbilder für die Schüler*innen darstellen und zudem implizit ihre Haltung gegenüber Ausgrenzungsverhalten vermitteln. Somit spielen sie eine wichtige Rolle im sozialen Gefüge ihrer Klasse (Mulvey et al., 2021).

Die vorliegende Studie untersucht anhand von hypothetischen Ausgrenzungsszenarien die Reaktionen von Lehrkräften und Lehramtsstudierenden auf soziale Ausgrenzung unter Schüler*innen in Abhängigkeit von verschiedenen Merkmalen des ausgeschlossenen Kindes. Darüber hinaus wird die Rolle der Empathie bei der Reaktion auf Ausgrenzung untersucht.

Methode

Es nahmen 355 Personen (44 männlich, 311 weiblich, M_{Alter}=31.36, SD=8.70) an der Studie teil, davon 287 Lehrkräfte im aktiven Schuldienst ($M_{Dienstjahre}$=7.31 Jahre, SD=6.36) und 69 Lehramtsstudierende aus verschiedenen in Deutschland gängigen Schulformen bzw. Lehramtsstudiengängen.

Den Teilnehmenden wurde ein Szenario präsentiert, in dem ein Kind im Physikunterricht durch andere Kinder von einer Gruppenarbeit ausgeschlossen wird. Die Teilnehmenden beurteilten das Ausgrenzungsverhalten (siebenstufige Skala, „nicht in Ordnung“ bis „in Ordnung“) und gaben an, wie wahrscheinlich es wäre, dass sie selbst in eine solche Situation eingreifen würden und für wie wahrscheinlich sie es halten, dass Kolleg*innen eingreifen würden (siebenstufige Skala, „sehr unwahrscheinlich“ bis „sehr wahrscheinlich“). Darüber hinaus wurde die dispositionelle Empathie (BES, Heynen et al., 2016) der Teilnehmenden erfasst.

In den Szenarien wurde folgende Merkmale des ausgeschlossenen Kindes variiert: Geschlecht (Jungenname vs. Mädchenname), Herkunft (typisch deutscher Name vs. typisch türkischer Name) und Sozialverhalten des Kindes (zusätzliche Beschreibung des Kindes mit typischen ADHS-Symptomen vs. keine zusätzliche Beschreibung). Daraus ergab sich 2x2x2 between-subjects Design, bei dem jeder (angehenden) Lehrkraft randomisiert eines der acht Szenarien zugewiesen wurde. Das Setting des Ausgrenzungsszenario war der MINT-Bereich, ein Feld, in dem Mädchen und Kinder mit Migrationshintergrund typischerweise unterrepräsentiert sind.

Ergebnisse

Mit einer univariaten ANOVA wurde untersucht, ob Merkmale des ausgeschlossenen Kindes (Geschlecht: männlich, weiblich; Herkunft: türkisch, deutsch; Sozialverhalten: Vorhandensein von ADHS-Symptomen, keine Erwähnung von ADHS-Symptomen) die Bewertung der Ausgrenzung durch die Lehrkräfte beeinflussen. Wie erwartet zeigte sich ein Haupteffekt des Sozialverhaltens, $F(1, 233)=5.85$, $p=.016$, $\eta_p^2=.02$, insofern, dass bei Kindern mit ADHS-Symptomatik die Ausgrenzung etwas weniger abgelehnt wurde ($M=2.51$, $SD=1.37$) als bei Kindern ohne ADHS-Symptomatik ($M=2.11$, $SD=1.33$). Zudem zeigte sich ein Interaktionseffekt aus Herkunft und Geschlecht des ausgeschlossenen Kindes, $F(1, 233)=4.54$, $p=.034$, $\eta_p^2=.02$: Bei türkischen Mädchen ($M=2.59$, $SD=1.52$) wurde die Ausgrenzung etwas weniger stark abgelehnt als bei deutschen Mädchen ($M=2.13$, $SD=1.34$). Bei Jungen zeigt sich dieser Herkunftseffekt nicht.

Zur Untersuchung der Interventionswahrscheinlichkeit wurde eine Messwiederholungs-ANOVA gerechnet mit den beiden Fragen zur Interventionswahrscheinlichkeit als Messwiederholungsfaktor und den Merkmalen des ausgeschlossenen Kindes (Geschlecht: männlich, weiblich; Herkunft: türkisch, deutsch; Sozialverhalten: Vorhandensein von ADHS-Symptomen, keine ADHS-Symptome) als unabhängigen Variablen. Hierbei ergab sich lediglich ein Haupteffekt des Messwiederholungsfaktors, $F(1, 232)=102.70$, $p<.001$, $\eta_p^2=.31$, der zeigt, dass Lehrkräfte bei sich selbst die Wahrscheinlichkeit höher einschätzen in die Situation einzugreifen als sie es von ihren Kolleg*innen erwarten würden. Die Merkmale des ausgeschlossenen Kindes hatten weder auf die eigene Interventionswahrscheinlichkeit einen Einfluss noch auf die erwartete Interventionswahrscheinlichkeit der Kolleg*innen.

Weiterhin zeigten Regressionsanalysen wie erwartet, dass Personen mit höher ausgeprägter Empathie das Ausgrenzungsverhalten stärker ablehnten ($b=-.73$, $t(222)=-3.84$, $p<.001$) und eine höhere Wahrscheinlichkeit angaben, in die Situation einzugreifen ($b=.53$, $t(222)=2.99$, $p=003$) als Personen mit geringerer Empathie.

Diskussion

Es lässt sich festhalten, dass Merkmale des ausgeschlossenen Kindes zwar relevant für die Bewertung der Ausgrenzungssituation waren, nicht aber die antizipierten Handlungstendenzen beeinflussten. Zudem scheint Empathie insgesamt eine wichtige Rolle bei der Reaktion von Lehrkräften auf soziale Ausgrenzung zu spielen und bietet somit einen Ansatzpunkt für Trainingsprogramme zum Umgang mit Ausgrenzungsverhalten unter Schüler*innen.

52. Unterschiede in der ethnischen Diskriminierungswahrnehmung im Ausbildungsübergang: Eine intersektionale Perspektive nach Herkunftsgruppen und Geschlecht

Julia Hufnagl
Otto-Friedrich-Universität Bamberg, Germany

Viele Jugendliche mit Migrationshintergrund erleben Chancengleichheit in Deutschland als unzureichend. Die berufliche Bildungsforschung fokussiert sich bisher insbesondere auf statistische Diskriminierung, während es hinsichtlich subjektiver Diskriminierungserfahrungen aus der Perspektive der Jugendlichen an empirischen Studien im Bildungskontext fehlt. Subjektive Erfahrungen stimmen nicht unbedingt mit den objektiv messbaren Ungleichheiten überein, sondern können stark davon abweichen. Für Betroffene ist die erlebte Situation meist wichtiger als die tatsächliche Diskriminierung.

Diese Studie stellt daher die Frage, wie sich die Wahrnehmung persönlicher ethnischer Diskriminierung bei Jugendlichen je nach Herkunftsgruppenzugehörigkeit unterscheidet. Mehrere Studien haben bereits Herkunftsgruppenunterschiede in der Wahrnehmung ethnischer Diskriminierung bestätigt. Die Besonderheit dieser Arbeit liegt darin, dass der Fokus auf dem Übergang zur beruflichen Bildung liegt. Die abhängige Variable ist dementsprechend die wahrgenommene ethnische Diskriminierung bei der Ausbildungsplatzsuche.

Das Geschlecht gilt neben dem Migrationsstatus als relevante soziale Kategorie in der Erforschung von Übergängen in die Berufsbildung. Um das Zusammenspiel der Herkunftsgruppe mit dem Geschlecht zu untersuchen, wird eine intersektionale Sichtweise eingenommen. Insbesondere im quantitativen Bereich fehlt es an intersektionalen Studien. Während Skrobanek zeigt, dass Frauen ethnische Diskriminierung stärker wahrnehmen als Männer, kommen andere Studien zu dem Ergebnis, dass Frauen weniger von ethnischer Diskriminierung berichten und ethnische Diskriminierung auf dem Arbeitsmarkt mit einer geringeren Wahrscheinlichkeit als Männer erwarten. Aufgrund der widersprüchlichen Ergebnisse stellt sich die Frage, wie sich die wahrgenommene ethnische Diskriminierung je nach Geschlecht unterscheidet. In welcher Herkunftsgruppe ist der Unterschied zwischen jungen Männern und Frauen am stärksten ausgeprägt?

Produktivitätstheoretische Annahmen lassen vermuten, dass türkischstämmige Jugendliche mehr Diskriminierung als Jugendliche anderer Herkunftsgruppen wahrnehmen, da Namen und/oder phänotypische Merkmale insbesondere bei ihnen sichtbar auf einen Migrationshintergrund hinweisen. Hinsichtlich der Interaktion der Herkunftsgruppenzugehörigkeit mit dem Geschlecht ergeben intersektionale Theorien gegenläufige Erwartungen: Die Subordinate Male Target Hypothesis lässt vermuten, dass Migranten ethnische Diskriminierung stärker wahrnehmen als Migrantinnen. Der

Additive Approach und der Intersectional-Inspired Approach jedoch begründen die Annahme, dass Migrantinnen mehr persönliche ethnische Diskriminierung wahrnehmen als Migranten.

Zur Hypothesenüberprüfung werden Personen mit Migrationshintergrund aus Welle 5 der Startkohorte 4 des NEPS herangezogen (N = 1.421). Für Zugewanderte aus der Türkei (N = 307), der ehemaligen Sowjetunion (N = 296), Polen (N = 131) und dem ehemaligen Jugoslawien (N = 141) liegt eine ausreichend große Stichprobe für die Analysen vor. Um den Herkunftsgruppeneffekt zu erklären, wird ein logistisches Regressionsmodell berechnet. Ein erweitertes logistisches Interaktionsmodell gibt Auskunft darüber, ob die Stärke des Zusammenhangs der Herkunftsgruppenzugehörigkeit nach Geschlecht variiert. Die Berücksichtigung von Interaktionseffekten gilt als eine Möglichkeit, intersektionale Fragestellungen quantitativ zu beantworten. Da Effektgrößen (wie die Logit-Koeffizienten bzw. Odds-Ratios) zwischen den Modellen nicht vergleichbar sind, werden im Nachgang an die Regressionsmodelle marginale Effekte (AMEs, MERs und Predictive Margins) geschätzt. Robustheitstests geben zudem Auskunft über eine Veränderung der Ergebnisse bei einer Variation von Herkunftsgruppeneinteilungen, Generationenstatus und weiteren Einschränkungen der Analysepopulation.

Die Ergebnisse zeigen, dass junge Menschen aus der Türkei bei der Suche nach einem Ausbildungsplatz mehr persönliche Diskriminierung wahrnehmen als polnische, ex-jugoslawische und ex-sowjetische Jugendliche. Die Wahrscheinlichkeit, dass polnische im Vergleich zu türkischen Jugendlichen ethnische Diskriminierung erfahren, ist am geringsten. Es konnte kein signifikanter Zusammenhang zwischen persönlicher Diskriminierung und dem Geschlecht festgestellt werden, was vermutlich auf die geringe Stichprobengröße zurückzuführen ist. Dennoch ist in der türkischen Gruppe ein höherer Geschlechterunterschied als in den anderen Herkunftsgruppen erkennbar. In der türkischen Gruppe ist die Wahrscheinlichkeit bei Frauen höher, bei ex-jugoslawischen Zugewanderten jedoch bei Männern.

Die Ergebnisse lassen besser verstehen, wie diskriminierendes Verhalten von Jugendlichen im Übergang wahrgenommen wird. Diskriminierungswahrnehmungen sind ein wichtiger Indikator für gesellschaftliche Schieflagen und gelten als zentrale Grundlage für die Planung von Antidiskriminierungsmaßnahmen. Weitere intersektionale Betrachtungen von wahrgenommener Diskriminierung sind notwendig, denn obwohl Diskriminierungen ein alltägliches Phänomen auf dem Ausbildungsmarkt sind, fehlen Daten über die Prozesse, die ihr zugrunde liegen.

53. HAND in HAND: Ein Lehrkräfte-Trainingsprogramm zur Förderung Sozio-Emotionaler Kompetenzen und Diversitätsbewusstsein

Mojca Rožman[1], Nina Roczen[2], Johannes Hartig[2], Valerie Fredericks[3], Ana Kozina[4], Magnus Oskarsson[5], Saša Puzić[6], Gina Tomé[7]

[1]IEA, Hamburg; [2]DIPF, Frankfurt; [3]University of Graz, Austria; [4]Educational Research Institute (ERI), Ljubljana, Slovenia; [5]Mid Sweden University (MIUN), Sundsvall/Östersund, Sweden; [6]Institute for Social Research in Zagreb (ISRZ), Zagreb, Croatia; [7]University of Lisbon, Portugal

Theoretischer Hintergrund

Lehrkräfte sind mit einer Vielzahl von Herausforderungen wie neuen Qualifikationsanforderungen, rasanten technologischen Entwicklungen und einer zunehmenden sozialen und kulturellen Vielfalt konfrontiert. Das EU Erasmus+-Projekt „*HAND IN HAND: Empowering teachers across Europe to*

deal with social, emotional and diversity related career challenges" unterstützt Lehrer*innen mit einem System aus Trainings und regelmäßiger Supervision, diesen Herausforderungen erfolgreich zu begegnen. Das Programm soll hierzu sozio-emotionale Kompetenzen sowie Diversitätsbewusstsein fördern.

Das HAND in HAND-Programm ist ein achtsamkeitsbasiertes Trainingsprogramm, das familientherapeutische Elemente, Elemente aus Trainings zu Diversitätsbewusstsein und Antidiskriminierung sowie aus familientherapeutischen Ansätzen integriert. Das Programm wird in Kroatien, Österreich, Portugal, Schweden und Slowenien durchgeführt und durch ein Team in Deutschland evaluiert.

Basierend auf der Konzeptionalisierung der Collaborative for Academic Social and Emotional Learning (CASEL, 2013) unterscheiden wir Selbstwahrnehmung und Selbstmanagement als emotionale Kompetenzen. *Selbstwahrnehmung* wird definiert als die Fähigkeit, die eigenen Emotionen, körperliche Empfindungen und Gedanken sowie deren Einfluss auf das eigene Verhalten zu erkennen. Unter *Selbstmanagement* wird die Fähigkeit verstanden, die eigenen Emotionen, körperliche Empfindungen, Gedanken und Verhaltensweisen in verschiedenen Situationen angemessen zu regulieren.

Soziale Kompetenzen setzen sich nach der Konzeptualisierung von CASEL zusammen aus Sozialem Bewusstsein und Beziehungskompetenz. *Soziales Bewusstsein* beschreibt die Fähigkeit, verschiedene Perspektiven einzunehmen und unterschiedliche Sichtweisen akzeptieren, Respekt gegenüber anderen sowie Wertschätzung von Diversität. *Beziehungskompetenz* bezieht sich auf die Fähigkeit, konstruktive Beziehungen aufzubauen und zu pflegen – ein zentraler Aspekt ist dabei die Kommunikationskompetenz (CASEL, 2013).

Diversitätsbewusstsein wird definiert als „a person's acknowledgement of culture and social context variables like class, race, ethnicity, gender, sexual orientation, physical ability, and religion" (Mosley-Howard, Witte, & Wang, 2011, p. 66). Diversitätsbewusstsein wird außerdem verstanden als eine Kombination von Wissen, Fähigkeiten und Einstellungen, die für das Leben in soziokulturell diversen Gesellschaften erforderlich sind. Diese sind Voraussetzung für das kritische Verständnis von Unterdrückung, Ungleichheiten und Ungerechtigkeit in der Gesellschaft sowie für das soziale Handeln.

Im Rahmen des Projekts soll experimentell untersucht werden, ob sozio-emotionale Kompetenzen und Diversitätsbewusstsein in der HAND in HAND Trainingsgruppe (Experimentalgruppe) im Zeitraum zwischen Prä- und Posttest in allen fünf Ländern (Kroatien, Österreich, Portugal, Schweden, Slowenien) signifikant gesteigert werden und ob die Zusammenarbeit der Lehrkräfte signifikant verbessert wird – jeweils im Vergleich zu der Kontrollgruppe, die kein HAND in HAND Training erhält.

Zur Präsentation dieses Beitrages liegen lediglich Daten zum Prätest vor. Anhand dieser Daten soll untersucht werden, ob die eingesetzten Messinstrumente ausreichend reliabel sind und die dimensionale Struktur den Erwartungen entspricht.

Methode

In jedem der fünf Länder wurden 12-20 Schulen (insg. *N* = 1000 Lehrkräfte sowie Schulleitungen und weiteres Schulpersonal) zur Teilnahme an der Interventionsstudie eingeladen. Anschließend wurden die Schulen nach dem Zufallsprinzip entweder der Experimentalgruppe (6-10 Schulen pro Land) oder der Kontrollgruppe (6-10 Schulen pro Land) zugewiesen.

Zur Vorher- und Nachher Messung von sozio-emotionalen Kompetenzen und Diversitätsbewusstsein wurden bestehende Selbstberichtsskalen zusammengestellt. Außerdem werden die Probanden nach Beendigung des Programms im Rahmen von Fokusgruppeninterviews befragt werden.

Die Dimensionalität der Skalen wird faktorenanalytisch untersucht, die Reliablität wird über die interne Konsistenz ermittelt. Außerdem wird die Messinvarianz zwischen Ländern und Experimentalgruppen anhand von Mehrgruppen-CFA überprüft. Die Effekte des Trainings auf sozio-emotionale Kompetenzen und Diversitätsbewusstsein sowie Lehrerkooperation werden varianzanalytisch untersucht.

Ergebnisse und ihre Bedeutung

Im Rahmen dieses Beitrages werden Ergebnisse zu den Prätest-Daten vorgestellt. Diese umfassen Informationen zur Reliabilität der Skalen, zur Messinvarianz sowie ihrer Struktur. Das Programm ist ein „Policy Experiment“ mit intensiver Beteiligung von Akteuren aus der Bildungspolitik, so dass im Falle einer positiven Evaluierung des Trainings Bedingungen für eine effektive, möglichst europaweite Dissemination geschaffen sind. Damit kann ein wichtiger Beitrag zur Unterstützung von Lehrkräften im konstruktiven Umgang mit Herausforderungen des Schulalltags, zur Verbesserung des Wohlbefindens von Lehrkräfte und Schüler*innen und zur Erreichung eines inklusiven Schulklimas geleistet werden.

54. Bereits im Studium Lehramtsstudierende auf den Umgang mit Belastungen in inklusiven Settings vorbereiten - Erste Evaluationsergebnisse eines Seminars zum ressourcenorientierten Umgang mit Belastungen

Victoria Wiederseiner
Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

In diesem Poster-Beitrag werden erste Evaluationsergebnisse von gesundheitsförderlichen und ressourcenorientierten Seminarbausteinen zum Umgang mit Belastungen in inklusiven Settings für Grundschullehramtsstudierende berichtet und vor dem Hintergrund deren Bedeutung für die erste Phase der Lehrkräfteprofessionalisierung diskutiert.

Empirische Untersuchungen zeigen, dass durch gesellschaftliche Transformationsprozesse (z.B. UN-Behindertenrechtskonvention, Fluchtmigration) und die damit verbundene steigende Heterogenität der Schülerinnen und Schüler bereits bei Lehramtsstudierenden negative Bewältigungsmuster der beruflichen Belastungen in inklusiven Settings in einem weiten Begriffsverständnis zu finden sind (Kerkhoff, 2019; Hohensee & Schiemann, 2021). Studien zur Gesundheit von (angehenden) Lehrpersonen betonen für ein individuelles Belastungsmanagement zudem das Verhältnis von Belastungen und Ressourcen (Bakker & Demerouti, 2007) sowie den Ausbau der ressourcenorientierten und professionellen Handlungskompetenzen, welche durch entsprechende Unterstützungsangebote gefördert werden können (Döring-Seipel & Dauber, 2010). Im Sinne der Lehrkräfteprofessionalisierung sollten nach den KMK-Kompetenzen bereits in der ersten Phase der

Ausbildung von Lehrkräften Lerngelegenheiten geschaffen werden, die sich sowohl mit der Auseinandersetzung von „Ergebnissen der Belastungs- und Stressforschung“ als auch mit dem „Umgang mit Belastungen“ befassen (KMK, 2019). Bislang lassen sich jedoch in der universitären Lehrkräftebildung kaum Lerninhalte zum Umgang mit der Bewältigung dieser beruflichen Belastungen im Sinne einer Förderung der gesundheitsbezogenen Kompetenzen der Lehramtsstudierenden im Rahmen eines nachhaltigen und langfristigen Bildungsprozesses in der Lehrkräfteprofessionalisierung finden (Lamanauskas, 2018).

Ziel der Seminarbausteine ist es daher, die persönlichen Ressourcen der teilnehmenden Studierenden im Zuge des Empowerment-Ansatzes (Kliche & Kröger, 2008) frühzeitig im Lehramtsstudium zu stärken. Demnach sollen die Studierenden selbst befähigt werden, ihre Ressourcen für den Umgang mit Belastungen in inklusiven Settings für ihren späteren Schulalltag zu nutzen. Die einzelnen Lerneinheiten orientieren sich dabei an der Fortbildungskonzeption „LehrKRÄFTE schonen und sinnvoll einsetzen“ (Elting et al., 2021), welches für die erste Phase der Lehrkräfteausbildung adaptiert und weiterentwickelt wurde. Folgenden Forschungsfragen soll dabei nachgegangen werden:

• Welche Belastungen und welche Ressourcen geben die Grundschullehramtsstudierenden vor und nach dem Seminar an, um mit Belastungen in inklusiven Settings umzugehen und wie verändern sich diese?

• Wie bewerten die Grundschullehramtsstudierenden den Lernnutzen und die Transferwirkung des Seminars?

Die Seminarbausteine wurden an der FAU Erlangen-Nürnberg im Wahlpflichtbereich in einer universitären Lehrveranstaltung mit Grundschullehramtsstudierenden (N = 34; Durchschnittsalter M = 21.10, SD = 3.32; 87% weiblich; Semester M = 3.40, SD = .84) und einer Vergleichsgruppe einer weiteren universitären Lehrveranstaltung zum professionellen Umgang mit Heterogenität (N = 32; Durchschnittsalter M = 23.01, SD = 6.10; 78% weiblich; Semester M = 4.30, SD=.79) durchgeführt und in einem Prä-Post-Design evaluiert. Die Studierenden wurden hierzu mittels Fragebögen zu Beginn und am Ende des Semesters unter anderem zu ihrem Belastungserleben und den Ressourcenbereichen (inklusionsbezogene Selbstwirksamkeit, Einstellung zu Inklusion und soziale Unterstützung) befragt. Zudem schätzten die Lehramtsstudierenden in einem vierstufigen Ratingformat (1 = „stimme gar nicht zu“ bis 4 = „stimme voll zu“) den Lernnutzen und die geplante Anwendbarkeit der Seminarinhalte ein. Die ersten Ergebnisse zeigen, dass die inklusionsbezogene Selbstwirksamkeit nach dem Besuch des Seminars von den Studierenden der Seminargruppe in Gegenüberstellung zu der Vergleichsgruppe signifikant positiver eingeschätzt wird ($t(26)=4.04$, $p<.05$, $d=.73$). Zudem weisen die anderen Ressourcenbereiche hinsichtlich der sozialen Unterstützung und der Einstellung zu Inklusion nach dem Besuch des Seminars in der Seminargruppe einen tendenziell positiven Anstieg zu Semesterende auf. Der gesamte Lernnutzen der Seminarbausteine (M=3.62; SD=.59) und die geplante Anwendbarkeit der Seminarinhalte (M=3.71; SD=.35) werden von den Studierenden insgesamt nahe am Maximalwert bewertet. Besonders die praxisorientierten Module 3 (Lernnutzen M= 3.58, SD= .39; Anwendbarkeit der Seminarinhalte M= 3.45, SD= .43) und 5 (Lernnutzen M= 3.55, SD= .41; Anwendbarkeit der Seminarinhalte M= 3.48, SD= .52) werden von den

Studierenden sehr hoch bewertet. Die Ergebnisse werden vor dem Hintergrund der Weiterentwicklung adäquater Bausteine für die Lehrkräfteprofessionalisierung zum Umgang mit Belastungen in inklusiven Settings diskutiert.

55. Unterstützung des Hörverstehens von Lernenden mit heterogenen Lernvoraussetzungen am Ende der Sekundarstufe I

Anke Schmitz[1], Hanne Brandt[1], Björn Rothstein[2]
[1]Leuphana Universität, Deutschland; [2]Ruhr Universität Bochum

Theoretischer Hintergrund

Das Hörverstehen ist eine fachübergreifende Sprachkompetenz (Hecker et al., 2015; KMK, 2022), die in Zeiten der fortschreitenden Digitalisierung besondere Aktualität erhält. Nicht zuletzt im Kontext des Distanzlernens wurden informierende Hörtexte von Lehrkräften in digitalen Lernumgebungen bereitgestellt (Wagner et al., 2020). Das Verstehen von Hörtexten stellt jedoch im Vergleich mit gedruckten Texten hohe Anforderungen an leistungsschwächere Lernende (Marx & Roick, 2012). Hörtexte sind nicht nur flüchtig, sie beinhalten auch bildungssprachliche Merkmale und Strukturen (Gogolin & Lange, 2010; Rubin et al., 2000) und weichen von gewohnten alltagssprachlichen Verstehenssituationen ab (Berendes et al., 2013; Marx et al., 2016). Deshalb kommt der sprachlichen Verständlichkeit für das Verstehen solch komplexerer Hörtexte ein zentraler Stellenwert zu (Kürschner & Schnotz, 2008), der bislang noch nicht systematisch untersucht wurde. Der vorliegende Beitrag untersucht, ob ausgewählte Prinzipien aus der Textverständlichkeitsforschung (Clinton-Lisell, 2022; Kürschner & Schnotz, 2008) das Hörverstehen von Lernenden unterstützen können und zur Verständlichkeit beitragen. Betrachtet werden globale strukturelle Kohärenzbildungshilfen, die den Hörtext auf globaler Textebene strukturieren (Arfè et al., 2018; Schnotz, 1994; Schmitz & Gräsel, 2016) und roten Faden beim Zuhören bieten.

Fragestellungen und Hypothesen

Untersucht wird, ob globale Kohärenzbildungshilfen das Hörverstehen unterstützen und ob Lernende mit unterschiedlichen Lernvoraussetzungen (thematisches Vorwissen; soziokulturelle Hintergrundmerkmale; Sprachkompetenz im Deutschen) wie bei Lesetexten (u.a. McNamara & Kintsch, 1996; Ozuru et al., 2009; Schmitz, 2017) differenziell profitieren. Erwartet wird, dass Lernende mit schlechteren Voraussetzungen stärker von den Kohärenzbildungshilfen profitieren als Lernende mit besseren Voraussetzungen. Denkbar ist bei Letzteren auch ein Umkehreffekt (McNamara & Kintsch, 1996), wonach Kohärenzbildungshilfen zu einer oberflächlicheren Rezeption und schlechteren Leistungen führen.

Methode

An der randomisierten kontrollierten Studie nahmen 153 Neuntklässler:innen (53 % 45% w; 2% d) an Gymnasien und Oberschulen teil. Ihnen wurde ein Hörtext mit einem erhöhten oder niedrigen globalen Kohärenzgrad vorgespielt. Die UV „globale Kohärenzbildungshilfe“ wurde in zwei Stufen operationalisiert: Ein erhöhter Kohärenzgrad (G+) wurde realisiert durch totale Rekurrenz (wörtliche Wiederholung von Begriffen), Hypophora in Form von reflexiven Fragen analog zu Zwischenüberschriften in Texten, Explikation semantischer Relationen (Nennung von semantischen Makrostrukturen wie Problem, Lösung). In dem wenig kohärenten Hörtext (G-) wurden diese

Kohärenzbildungshilfen weggelassen. Inhaltlicher Gegenstand war ein curricular relevantes Thema aus dem Deutschunterricht. AV war das Verstehen des erklärten Sachverhalts, das mittels schriftlichem Test (21 Items, α = .79) erhoben wurde. Die MC- und halb-offenen Items erfassten wörtliches Verstehen, textbezogenes Schlussfolgern und das Verstehen impliziter Bedeutungen (Hecker et al., 2015). Weiter wurden das thematische Vorwissen (18 Items, α = .76), soziokulturelle Aspekte (Marx & Roick, 2012; Müller & Stanat, 2006) und Sprachfähigkeiten im Deutschen (C-Test, 30 Items, α = .93) erhoben.

Erste Ergebnisse und ihre Bedeutung

Die Lernenden erzielten mit dem wenig kohärenteren Hörtext im Durchschnitt M = 6.78 (SD = 4.11) Punkte und mit dem kohärenten Hörtext M = 8.00 (SD = 3.96) Punkte. Der Gruppenunterschied ist zum 5 %-Niveau signifikant (t(151) = -1.83, p = .03, d = -.30). Regressionsanalysen zeigen, dass der Effekt der Kohärenzbildungshilfen unter Kontrolle des thematischen Vorwissens und der sprachlichen Fähigkeiten im Deutschen (C-Test) signifikant bleibt (F(3 149) = 29.26, p = .000, adj. R2 = .36). Bei Schüler:innen, die kein Gymnasium besuchten (N = 62), ist der Effekt der Kohärenzbildungshilfen stärker ausgeprägt (part. η2 = .16) als bei den Gymnasiast:innen (N = 91, part. η2 = .06). Jedoch können die Lernenden nicht zu den Gymnasiastinnen aufschließen. Weitere differenzielle Analysen, in denen soziokulturelle Hintergrundmerkmale der Schüler:innen berücksichtigt werden, liegen zum Zeitpunkt der Tagung vor. Die Studie zeigt aber bereits jetzt, dass Verständlichkeitsmerkmale von Lesetexten auch das Verstehen von Hörtexten begünstigen und insbesondere leistungsschwächere Schüler:innen profitieren. Der positive Effekt hat insofern Implikationen, inwiefern Lehrkräfte ihre Lernenden mittels Scaffolds gezielt beim Verstehen von Hörtexten unterstützen können.

56. A longitudinal study of gender differences in the benefits of elementary school students from digitized math trainings.

Noro Schlorke
TU Dortmund, Germany

Gender differences in Numerical Fluency have been the subject of several studies, and these studies revealed contradicting results. Whereas no notable gender difference in mathematical abilities were found in elementary school children (Aragón, Serrano, & Navarro, 2018; Hyde, 2005; Lachance & Mazzocco, 2006; Pina et al., 2021), boys performed better on certain mathematical tests with increasing age (Hyde et al., 1990; Mullis, Martin, Fierros, Goldberg, & Stemler, 2000; Roselli, Ardila, Matute, & Inozemtseva, 2009). An advantage in the performance of boys was recorded especially in tests of spatial math (Roselli et al., 2009; Liu & Wilson, 2009). Interestingly, boys begin to gain an advantage over girls as both genders develop different attitudes towards computer games (Ziemek, 2006). Together with a higher interest in computer games, boys develop a higher self-esteem and a more positive attitude towards computers (Meelissen, Drent 2008; Fraillon, Ainley, Schulz, Friedman, Gebhardt 2014) and use them more for recreational purposes than girls do, while girls use computers more for communication and academic achievement (Johnson, 2011; Fraillon et al., 2014). The gender difference in attitude towards computers might change with age, as Christensen et al. (2005) reported a higher enjoyment of computer usage in fourth and fifth grade girls than in younger girls. Boys' more positive attitudes towards computers as well as biases of teachers might cause boys'

higher enjoyment and literacy in computer use as compared to girls (ACARA, 2015). These differences indicate the potential challenge to the digitalization of math education to avoid that pre-existing gender differences in favour of boys might even increase during interventions. The present study evaluates one such intervention with the focus on its differing effect on boys and girls. This longitudinal study aims to investigate the effects of an education software on numerical fluency. Specifically, we investigated whether the learning software effects on students with deficits in numerical fluency found in previous studies would also prevail in regular elementary school students. As there are gender differences in numerical fluency and computer use, gender might be a moderator of the intervention effect. We conducted a field experiment with a single treatment group and a control group. Two subsequent cohorts of students of one elementary school were tested in classes. The second-grade classes were assigned to the treatment group, while two classes of third graders were assigned to the control group. The second grade students used a tablet with the “Blitzrechnen”-App (Klett) three times a week for at least 15 minutes each starting June 2017. The training began in second grade and was conducted until the end of the first term of fourth grade. In line with the hypothesis , the average treatment effect on numerical fluency was statistically significant (ATE = 13.6, $p < .001$, see Figure 1). However, subsequent analysis demonstrated that the treatment effect for girls was not significant (ATEgirls= 4.0, $p = .56$), whereas the effect for boys was (ATEboys= 20.4, $p < .001$). These results support the assumptions raised in our research question. Testing the interaction by means of the procedure proposed by Robinson et al. (2013), the difference slightly missed significance (bdiff = -.19, $p = .074$), showing at least a tendency that the training was more effective for boys than for girls.

57. Warum geht Inklusion nicht (für alle)? Zum Zusammenhang von Lerntheorie, Behinderungsverständnis und dem gemeinsamen Unterricht

Tom Selisko, Christine Eckert, Franziska Perels
Universität des Saarlandes, Deutschland

Ziel

Die Beschulung von Kindern mit Behinderung wird bevorzugt im Rahmen der Inklusion diskutiert (z. B. Avramidis et al., 2019; Forlin et al., 2011; Saloviita, 2020). Die Inklusion steht allerdings vor der Herausforderung, dass ihre Definition selbst anhaltender Gegenstand des Diskurses ist (Grosche & Vock, 2018). Ergänzend zur Einstellung zur Inklusion (bzw. Exklusion) sind zwei weitere Dimensionen für die Einstellung zur Beschulung von Kindern mit Behinderung bedeutsam: die Vorstellung vom Lehren und Lernen (Södervik et al., 2022) und das zu Grunde liegende Behinderungsmodell (Gebhardt et al., 2022). Diese schaffen die Grundlage dafür, in welchem Ausmaß und ob überhaupt eine inklusive Beschulung als möglich und sinnvoll erachtet werden kann. Das Ziel der Arbeit ist die Entwicklung eines Instruments zur Erfassung der Einstellung zur Beschulung von Kindern mit Behinderung.

Theoretischer Hintergrund

Was als Behinderung wahrgenommen wird, ist dabei abhängig vom zu Grunde liegenden Behinderungsverständnis. Im Rahmen der vorliegenden Untersuchung werden drei Modelle von Behinderung unterschieden: Das medizinische Modell, welches Behinderung als objektive und

messbare Funktionseinschränkung definiert (Shakespeare, 2017), das soziale Modell, welches Behinderung als Teilhabeeinschränkung auf Grund gesellschaftlicher Barrieren definiert (Oliver & Barnes, 2010), und das relationale Modell, welches die Wechselwirkungen zwischen den Aspekten eines medizinischen und eines sozialen Modells fokussiert (Waldschmidt, 2020). Während also im medizinischen Modell die Behinderung im Individuum verortet ist, ist sie im sozialen Modell eine veränderbare Eigenschaft der Umwelt.

Die Heterogenität im Klassenzimmer in Regelschulen nimmt mit einer wachsenden Zahl von Kindern mit Behinderung zu (Lübke et al., 2016). Ob dies als grundsätzlich problematisch wahrgenommen wird, ist insbesondere abhängig von den Lehr-Lernüberzeugungen der Lehrkraft: Konstruktivistische Überzeugungen gehen grundsätzlich von einem individuellen und selbstgesteuerten Lernprozess aus und sind somit auf einen individualisierten Unterricht ausgerichtet, während bei transmissiven Überzeugungen die steigende Heterogenität eine Mehrbelastung darstellt, da Inhalte immer weiter differenziert werden müssen (Södervik et al., 2022). Dementsprechend wird auf der Grundlage dieser theoretischen Vorüberlegungen ein Zusammenhang zwischen konstruktivistischen Lehr-Lernüberzeugungen, einem sozialen Verständnis von Behinderung und einer positiven Einstellung zur gemeinsamen Beschulung von Kindern mit und ohne Behinderung ausgegangen; gleichzeitig von einem Zusammenhang zwischen einem transmissiven Verständnis von Behinderung, einem medizinischen Modell von Behinderung und einer negativen Einstellung gegenüber der gemeinsamen Beschulung.

Die vorliegende Arbeit beschäftigt sich mit der Integration dieser drei Aspekte in ein Instrument zur ganzheitlichen Erfassung der Einstellung zur Beschulung von Kindern mit Behinderung.

Methode

Zur Erfassung des Behinderungsverständnisses wurde die Skala von Gebhardt et al. (2022) zu Grunde gelegt und durch eine selbsterstellte Skala zur Erfassung eines relationalen Modells ergänzt. Die Lehr-/Lernüberzeugungen wurden mit Hilfe der Skalen der BilWiss-Studie zur Erfassung konstruktivistischer und transmissiver Überzeugungen über das Lernen erhoben (Kunter et al., 2020). In Anlehnung an die Teacher's Attitudes Towards Inclusion Scale (Boyle et al., 2013) wurden darüber hinaus Items erstellt, die die Zustimmung oder Ablehnung zur gemeinsamen Beschulung von Kindern mit und ohne Behinderung erfragen. Die Likert-Skalen wurden zwecks besserer Vergleichbarkeit einheitlich gestuft. In einer Pilotierungsstudie mit N=12 Teilnehmenden (Lehramtsstudierende, vier männlich, acht weiblich, im Alter zwischen 19 Jahren und 27 Jahren) wurde der Fragebogen im Paper-Pencil- Format in einem ersten Schritt unter Verwendung von R(-Studio) evaluiert.

Ergebnisse

Die verwendeten Skalen erreichten eine zufriedenstellende Reliabilität mit einem Cronbachs Alpha im Bereich zwischen .60 und .81. Nach Analyse der Item-Trennschärfe und entsprechendem Ausschluss einzelner Items konnten Cronbachs Alpha-Werte im Bereich zwischen .71 und .88 festgestellt werden. Dabei korrelierten die Konstrukte erwartungsgemäß in kleinem bis mittleren Umfang (r zwischen .14 und .32). Unter Beachtung der in der Pilotstudie gewonnen Erkenntnisse wird das Instrument im nächsten Schritt als Onlinebefragung auf eine angemessen große Stichprobe angewendet, die die

Durchführung einer konfirmatorischen Faktorenanalyse erlaubt. Diese Ergebnisse werden im Rahmen der Tagung berichtet.

58. Informationen im Internet bewerten. Entwicklung und empirische Evaluation eines Messinstruments für Grundschulkinder

Tina Jocham, Sanna Pohlmann-Rother
Universität Würzburg, Deutschland

1 Theoretischer und empirischer Hintergrund

Der Zugang zu Informationen wird durch die Digitalisierung erleichtert. Unverändert bleibt die Kompetenz, sich mit diesen Informationen kritisch auseinanderzusetzen (Çetta et al., 2020). Die Informationsbewertung gilt deshalb als wichtige Teilkompetenz für schulisches und außerschulisches Lernen sowie die gesellschaftliche Teilhabe (Bos et al., 2014), die entwickelt werden muss, um ein Zurechtfinden in allen Lebensbereichen zu ermöglichen (Freimanis et al., 2013). Deutlich wird dies im Umgang mit Fake-News und Deepfakes in digitalen Kommunikations- und Informationskanälen (z.B. Gabriel, 2018). Dass dies nicht nur Erwachsene oder Schüler:innen der Sekundarstufe I und II betrifft, zeigt sich darin, dass Kinder immer früher Rezipienten digitaler Medien sind. Betrachtet man die Nutzungshäufigkeit von Kindern zwischen 6 und 13, fällt auf, dass zwei Drittel der 8- bis 9-jährigen Kinder regelmäßig online sind. Dabei bilden WhatsApp, Google, das Anschauen von Filmen und Videos sowie die Nutzung von YouTube die Top Vier der Internettätigkeiten (mpfs, 2020), wobei WhatsApp, YouTube und TikTok am beliebtesten sind (Reppert-Bismarck, et al., 2019). Die Internetpräferenzen der Kinder liegen damit im Wesentlichen in außerschulischen Bereichen und nicht auf der Verwendung von kindgerechten Angeboten, die häufig in schulischen Kontexten thematisiert werden (z.B. Blinde Kuh, fragFINN). Forschungsprojekte konzentrieren sich meist auf den Rechercheprozess im Internet mit schulischen Aufgaben (Gerhardts, 2020; Feil et al., 2013), während quantitative Erhebungen regelmäßig das Nutzungsverhalten der Kinder in den Blick nehmen. Im Mittelpunkt stehen dabei insbesondere die Nutzungshäufigkeit und -dauer, die Medienart und elterliche Kontrolle (z.B. KIM-Studie) sowie die Erfassung computer- und informationsbezogener Kompetenzen in der Sekundarstufe (Eickelmann et al. 2019) oder bei Erwachsenen (Meßmer et al., 2021). Zur Modellierung und standardisierten Erfassung von Informationskompetenzen im Grundschulbereich liegen demgegenüber kaum Erkenntnisse vor. Diese wären jedoch nötig, um spezifische Interventionen zu entwickeln und diese zu evaluieren (Hermida, Hielscher & Petko, 2017). Dass hierfür ein Bedürfnis besteht, zeigt insbesondere der Umstand, dass sich 64 % der 9- bis 11-jährigen Kinder zutrauen, Suchbegriffe im Internet zu verwenden, sich aber nur 14 % auch in der Lage fühlen, die gefundenen Informationen auf ihren Wahrheitsgehalt zu überprüfen (Hasebrink, et al., 2019). Inwiefern Grundschulkinder tatsächlich in der Lage sind, Informationen im Internet zu bewerten und welche Variablen hierauf Einfluss nehmen, ist derzeit empirisch nicht erfasst.

2 Fragestellung und methodisches Vorgehen

Das Ziel der Studie ist es, ein Modell für den Teilbereich der Informationsbewertung zu entwickeln und empirisch zu überprüfen, das kognitive, motivationale und affektive Variablen sowie internet- und quellenspezifische Merkmale integriert. Der Einfluss motivationaler Faktoren ist beispielsweise ebenso wenig bekannt wie das Themeninteresse der Schüler*innen oder die wahrgenommene

Relevanz der Aufgabe (Lewandowski & Kammerer, 2021). Dabei stellt sich die Frage, welche Variablen besonders auf die Informationsbewertung im Internet einwirken und wie diese zusammenhängen. Hierfür wurden zunächst Kompetenzen und Einflussfaktoren der Informationsbewertung deduktiv abgeleitet. Mithilfe explorativer Expert:innen-Interviews konnten weitere Faktoren für das Modell induktiv ergänzt werden.

Basierend auf dem entwickelten Modell entstand ein digitales *Messinstrument* für Grundschulkinder der dritten und vierten Jahrgangsklasse, das empirisch überprüft werden soll. Zur Erhöhung der Inhaltsvalidität beinhaltet das Instrument multimodale Testaufgaben aus der Lebenswelt der Kinder (z.B. YouTube-Videos, Screenshots), welche anhand von altersgerechten Kriterien eingeschätzt werden sollen. Da indes kaum fundierte, spezifische Kriterien für den Grundschulbereich vorliegen, wurden mögliche Lösungsstrategien auf Basis empirischer Studien (z.B. erhöhter Fokus auf gesponserte Inhalte oder hohes Vertrauen in Internetinhalte) abgeleitet. Die Relevanz dieser metakognitiven Strategien zur Lösung der Testaufgaben werden durch Expert:innen vorab überprüft und angepasst. Auf diese Weise soll untersucht werden, welche Bewertungskriterien Kinder in dieser Altersgruppe kennen und anwenden können (z.B. Intentionalität von Informationen).

3 Ergebnisse

Die Datenerhebung startet im Herbst 2022, sodass erste Ergebnisse im Rahmen des Einzelbeitrags vorgestellt werden.

59. Erprobung eines Videospieltrainings für räumliche Fähigkeiten im Ingenieursstudium

Nils Nolte[1], Seyedahmad Rahimi[2], Valerie J. Shute[3], Detlev Leutner[1]
[1]Universität Duisburg-Essen, Deutschland; [2]University of Florida, Gainesville, FL, US; [3]Florida State University, Tallahassee, FL, US

Räumliche Fähigkeiten stehen in besonderem Zusammenhang mit Erfolg in MINT-Fächern (Wai et al., 2009). Verschiedene Interventionsansätze wurden in der Vergangenheit erprobt, um räumliche Fähigkeiten zu trainieren und durch Transfereffekte unterschiedliche Erfolgskriterien in verschiedenen MINT-Fächern zu verbessern. Neben direkten Trainings räumlicher Fähigkeiten durch entsprechende Aufgaben, die oft auch für Tests verwendet werden (bspw. Cheng & Mix, 2014; Sorby et al., 2018), ist die Nutzung räumlich anspruchsvoller Videospiele (Shute et al., 2015), die die räumlichen Fähigkeiten der Spielenden ansprechen, ein vielversprechender indirekter Trainingsansatz.. Videospiele sind sehr verbreitet und beliebt (430 Mio. Videospielende in Europa, Newzoo, 2022; 97% US-amerikanischer Jugendlicher, Geiger, 2008), sodass mit höherer Teilnahmemotivation zu rechnen ist als bei anderen Trainingsansätzen.

Ziel der vorliegenden Studie war, Bauingenieur-Erstsemester ein Videospiel spielen zu lassen um ihre räumlichen Fähigkeiten und indirekt auch ihren Fachwissenserwerb und ihre Klausurnoten im Fach Technische Mechanik zu verbessern. Eingebettet in einen optionalen Kurs wurden 86 Bauingenieur-Erstsemesterstudierende teilrandomisiert einer von zwei Interventionsgruppen zugeteilt: Die Experimentalgruppe erhielt das räumlich anspruchsvolle Puzzlespiel “Portal 2”, während die Kontrollgruppe das Rätselspiel “Sherlock Holmes jagt Jack the Ripper” erhielt. Beide Gruppen sollten ihr jeweiliges Spiel innerhalb von vier Wochen jeweils acht Stunden lang spielen. Vor der Trainingsperiode, im direkten Anschluss sowie am Semesterende wurden die räumlichen Fähigkeiten

(mentaler Rotationstest, Vandenberg & Kuse, 1978), logisches Schlussfolgern (Subtests 2 und 3 des Culture Fair Tests, Weiß, 2006) sowie das Fachwissen in Technischer Mechanik (Dammann, 2016) erhoben. Aus der Portal 2-Gruppe haben 42 Studierende (von 54) das Training abgeschlossen und alle Tests bearbeitet, in der Sherlock Holmes-Gruppe 26 (von 32). Abschließend wurden die Klausurnoten in Technischer Mechanik von den Lehrenden erfragt (26 Noten in der Portal 2-Gruppe, 18 Noten in der Sherlock Holmes-Gruppe).

Der Interventionscheck zeigte einen signifikanten Zuwachs räumlicher Fähigkeiten in der Portal 2-Gruppe (Delta-M = 0.60, $t(41) = 3.64$, p(einseitig) < .001, 95% CI [0.32, Inf], d = 0.56), nicht aber in der Sherlock Holmes-Gruppe (DM = 0.30, $t(24) = 1.39$, p(einseitig) = .176, 95% CI [-0.07, Inf], d = 0.28). Allerdings wurde der Unterschied im Zuwachs zwischen den beiden Gruppen nicht signifikant, (Delta-$M_{Zuwachs}$ = 0.30, $t(50.13) = 1.10$, p(zweiseitig) = .277, 95% CI [-0.24, 0.84]). Demzufolge lässt sich abschließend nicht sagen, ob das Spielen von Portal 2 die räumlichen Fähigkeiten tatsächlich besser trainiert als das Spielen von Sherlock Holmes. Allerdings zeigte sich in einer hilfsweise hinzugezogenen vergleichbaren Stichprobe aus früheren Erhebungen ohne Intervention keinerlei Verbesserung der räumlichen Fähigkeiten (Delta-M = -0.43, $t(68) = -2.88$, p(einseitig) = .997, 95% CI [-0.68, Inf]), sodass eher von einem wegen der geringen Stichprobengröße nicht eindeutigen Trainingseffekt in beiden Interventionsgruppen als von einem reinen Übungseffekt auszugehen ist.

In Hinblick auf die Studienerfolgsmaße ergaben sich signifikante Zusammenhänge der räumlichen Fähigkeiten der Studierenden nach dem Training mit sowohl dem Fachwissen am Ende des Semesters (b = 0.26, β = .50, SE = 0.05, p < .001) als auch mit dem Klausurerfolg (b = 0.05, β = .38, SE = 0.02, p = .006). Die Multiplikation des durchschnittlichen Zuwachses der räumlichen Fähigkeiten pro Gruppe mit dem Regressionskoeffizienten für die Vorhersage des Klausurerfolgs ergibt die zu erwartende Verbesserung des Klausurerfolgs durch das Training: In der Portal 2-Gruppe ergab sich eine durchschnittliche Verbesserung von 3 Prozentpunkten (etwa eine Drittelnotenstufe) bzw. 1.5 Prozentpunkte in der Sherlock Holmes-Gruppe.

Insgesamt deuten unsere Ergebnisse auf das Potenzial videospielbasierter Trainings räumlicher Fähigkeiten für Studienerfolg im Bauingenieursstudium hin. Allerdings sind weitere Studien notwendig, die einerseits größere, ausgewogenere Stichproben für höhere statistische Power und andererseits eine aktive Kontrollgruppe verwenden, die – anders als in der vorliegenden Studie – selbst keinen Effekt aufweist. Das hier verwendete Spiel „Sherlock Holmes“ könnte aufgrund seiner dreidimensionalen Spielwelt vielleicht ebenfalls Trainingseffekte gehabt haben.

60. Triangulativer Ansatz zur Optimierung und Evaluation des digitalen Escape Rooms „Escape Climate Change“

Silvia Fränkel[1], Sebastian Becker-Genschow[1], Nadine Großmann[2]
[1]Universität zu Köln, Germany; [2]Universität Bielefeld, Germany

Theoretischer Hintergrund und Forschungsstand

Escape Rooms (ER) stellen ein neues Konzept in der Gestaltung von Lernumgebungen dar (Fotaris & Mastoras, 2020). Grundlegende Idee der ER ist das kooperative Lösen von Rätseln, um aus einem Raum zu entkommen (Fotaris & Mastoras, 2020; Veldkamp et al., 2020). Hierbei werden hands-on und minds-on-Aktivitäten miteinander kombiniert und situiertes Lernen durch authentische Kontexte

ermöglicht (Veldkamp et al., 2020). Die Anforderung der Aufgaben im ER können dabei an die Fähigkeiten der Teilnehmenden angepasst werden (Fotaris & Mastoras, 2020). Wird ein Gleichgewicht zwischen Anforderungen und Fähigkeiten hergestellt, kann dies die Entstehung selbstbestimmter Motivationsqualitäten begünstigen (Csikszentmihalyi, 2019; Ryan & Deci, 2017) sowie die Wahrnehmung eigener Kompetenz und Selbstwirksamkeit bei einer erfolgreichen Lösung der Aufgabe hervorrufen (Ryan & Deci, 2017). Als Vorteile der Implementation eines ER werden die Förderung von Zusammenarbeit, sozialer Interaktion und Kommunikation sowie des kritischen Denkens, Problemlösens und der Kreativität der Teilnehmenden beschrieben (Fotaris & Mastoras, 2020; Großmann & Fränkel, *eingereicht*). Veldkamp et al. (2020) berichten in ihrem systematischen Review positive Auswirkungen auf das Vergnügen und das Engagement der Teilnehmenden, verweisen jedoch im Hinblick auf die Belohnungen und Zeitbeschränkungen eines ER auch auf die Entstehung extrinsischer Motivation (s. auch Ryan & Deci, 2017).

Zurzeit ist eine zunehmende Anzahl an analogen und digitalen ER-Konzepten für den Unterricht feststellbar (Veldkamp et al., 2020). Für die Wirksamkeit dieser Konzepte fehlt jedoch häufig die empirische Evidenz (Veldkamp et al., 2020). Veldkamp et al. (2020) fordern daher folgerichtig, dass bestehende ER-Konzepte eine empirische Evaluation benötigen, während zugleich Richtlinien zur Entwicklung und Implementation neuer Konzepte aufgestellt werden sollten. Neben einer zumeist fehlenden empirischen Evidenz ist eine theoretische Fundierung notwendig. Bisher werden ER eher vage im game-based learning verankert (Veldkamp et al., 2020).

Ziel des vorliegenden Projekts ist die Evaluation eines bestehenden digitalen ER (*Escape Climate Change*), der von den Unternehmen ECOMOVE International e.V. und GEO°BOUND UG im Rahmen eines Projekts des Bundesministeriums für Umwelt, Naturschutz und nukleare Sicherheit entwickelt wurde und aktuelle gesellschaftliche Herausforderungen (Klimawandel) aufgreift (ECOMOVE International e.V., 2020).

Fragestellungen

Wie ist das emotional-soziale und motivationale Erleben während der Bearbeitung des digitalen ER ausgeprägt?

Ist der digitale ER lernwirksam bezüglich Fachwissen?

Welche Hindernisse im fachlichen und sozialen Lernprozess treten bei der Bearbeitung des digitalen ER auf?

Gibt es Unterschiede zwischen einer Implementation des digitalen ER, des analogen ER und einer vergleichbaren Unterrichtseinheit ohne ER hinsichtlich des emotional-sozialen, selbstbestimmt motivierten und fachlichen Lernens?

Methode

Schüler:innen der Sekundarstufe I von Gymnasien, Gesamt- und Realschulen sollen an der vorliegenden Studie teilnehmen. Zur Optimierung und Evaluation des ER sollen in einer ersten Phase in einer Triangulation quantitative und qualitative Methoden kombiniert werden. Motivationale und kognitive Variablen (bspw. wahrgenommene Kompetenz und intrinsische Motivation, Wilde et al.,

2009; Wissenszuwachs) sollen quantitativ erfasst werden. In der qualitativen Studie sollen während der Durchführung des ER die Blickbewegungen der Schüler:innen mittels mobilem Eye-Tracking erfasst und im Anschluss Interviews zum emotional-sozialen und motivationalen Erleben und möglichen Schwierigkeiten während der Durchführung geführt werden. Anhand der erhobenen Daten soll der ER im Sinne des design-based-research (McKenney & Reeves, 2012) optimiert werden. Bei dieser Optimierung könnte eine Community of Practice bestehend aus schulischen, außerschulischen und universitären Akteur:innen einbezogen werden. In der zweiten Phase soll der optimierte ER mit einem quasi-experimentellen Design evaluiert werden. Hierzu sollen der digitale ER, eine analoge Version des ER sowie die Erarbeitung des Themas Klimawandel ohne die Implementation eines ER mit dem gleichen triangulativen Ansatz verglichen werden. Nach der Diskussion der Befunde, ggf. wieder in der Community of Practice, sollen schlussendlich Empfehlungen zur unterrichtlichen Implementation des ER in einer Handreichung formuliert werden. Falls Ergebnisse aus der ersten Phase bereits vorliegen, können diese neben dem Projekt auf der Tagung diskutiert werden.

61. No Silver Bullet: Development of a Prompt Design for Vocational Education Students

Herbert Thomann, Viola Deutscher
University Mannheim, Germany

Theoretical Background

In the near future, the demand for skilled labor will increase while that for routine activities will decrease (Bughin et al., 2018). The remaining work tasks, then, will be more complex, and domain-specific problem-solving skills will become more critical. In this regard, digital simulation-based learning environments can help to meet these changed circumstances (Rausch et al., 2021). Since these more immersive environments offer a more complex learning environment, educational prompts could be a suitable tool to support learners with difficulties. However, prompting in the context of vocational education has received little attention. Even recently, few researchers have devoted themselves to this topic (see Chen et al., 2021; Eshuis et al., 2022; van der Meij & de Jong, 2011).

Research Question

This study aims to investigate prompts to support vocational education students in a digital simulation-based learning environment, thus improving their learning achievement.

Methods

As part of the research project "Problem Solving Analytics in Office Simulations" (PSA-Sim), the web-based learning platform LUCA is being developed in which learners are provided with an office simulation in which they process an authentic work scenario on the topic of “supplier selection.” An experimental design was realized in which industrial management assistants and office administrators (age range = 17 – 36 years) of all three years of their apprenticeship from 34 classes divided over 21 schools in Baden-Württemberg worked for approx. 65 minutes in LUCA with a questionnaire following. They were randomly assigned either to the experimental group with educational prompts (n = 319) or the control group without prompts (n = 361).

A log data-based real-time analysis enables predefined assistance to be triggered (prompting) when learners exhibit specific behavior (Rausch et al., 2021). 27 different cognitive, metacognitive, and non-

cognitive learning prompts were embedded in this working scenario (Table 1). The cognitive prompts were tailored to learners' behavior (log data) as mistakes in the learning environment. For this purpose, an analysis of possible errors was conducted before the design of prompts. After that, a learning prompt and trigger conditions were designed for each possible mistake. For example, if learners choose an incorrect exchange rate when calculating, they receive a personalized prompt indicating the potential source of the error (Table 1; Nr. 1). Furthermore, so-called events can be used to individualize task processing. For example, personalized prompts can be triggered based on explicit information provided by the learners (Figure 1).

Results

Paired t-test results showed that there was no significant difference of the mean final scores of the experimental group (M = 43.97, SD = 17.77) and the control group (M = 45.92, SD = 17.36); t(678) = 1.4438, p < 0.1493, suggesting that prompts were not able to improve learners' scores. The data analysis and evaluation are ongoing and will be ready for the poster presentation.

High cognitive load due to the immersive learning simulation and the complex working scenario could be why prompts did not lead to the expected positive learning outcomes (e.g., Eckhardt et al., 2013). Log data also showed that learners often did not perceive the prompts because they were sent through the mail system and did not appear as pop-up windows. Furthermore, learners can feel hindered and disturbed if prompts contain too much information and additional calls for action. In this case, prompts are perceived as extra tasks outside the core learning environment. The prompt design was revised based on these findings and the questionnaire results. The data from an additional 200 vocational students are currently being collected to test the adapted prompt design.

62. Effekte kooperativer Rezeption von Erklärvideos auf den Erwerb von Wissen und Kompetenzen

Laura Schultze, Raphael Koßmann, Nicoletta Bürger, Melanie Fabel-Lamla, Christof Wecker
Universität Hildesheim, Deutschland

Erklärvideos als Form informellen Lernens sind längst Bestandteil der Lebenswelt vieler Jugendlicher. Anders als bei Erklärungen durch die Lehrkraft im Unterricht können Schülerinnen und Schüler sich Erklärungen in Videos so oft anschauen, bis sie die Inhalte verstanden haben und anwenden können. Insbesondere unter den Bedingungen der COVID-19-Pandemie boten Erklärvideos Lehrkräften die Möglichkeit, Erklärungen individuell durch die Schülerinnen und Schüler zuhause rezipieren zu lassen. Dieser Einsatz von Erklärvideos weist Überschneidungen mit bereits seit längerem untersuchten Flipped-Classroom-Formaten auf, bei denen die Schülerinnen und Schüler Erklärvideos in Vorbereitung auf den Unterricht zumeist individuell rezipieren (z.B. Frei et al. 2020; Wagner et al. 2020). Zur Frage, wie Erklärvideos als Bestandteil von Präsenzunterricht didaktisch sinnvoll eingesetzt werden können, ist dagegen bislang noch kaum systematisch geforscht worden (Fiorella 2021). Das individuelle Rezipieren von Erklärvideos könnte leicht zu einer sogenannten „Verstehensillusion“ (Kulgemeyer 2022, S. 125) führen, bei der die Rezipierenden überzeugt sind, die Inhalte ausreichend erfasst zu haben, ohne sie wirklich verstanden zu haben. Um solchen Tendenzen entgegenzuwirken, legt die Forschung zum multimedialen Lernen vor allem einen Fokus auf die Gestaltung der Videos, um die kognitive Belastung zu reduzieren und produktive Lernaktivitäten zu

begünstigen (Mayer & Fiorella 2021). Doch auch die Art der Einbindung von Erklärvideos in das Interaktionsgeschehen im Unterricht könnte dazu beitragen, dass produktive Lernaktivitäten bei Schülerinnen und Schülern angestoßen werden. So könnte sich etwa im Präsenzunterricht eine kooperative Form der Rezeption von Erklärvideos anbieten, bei der durch die Interaktion über die Inhalte von Erklärvideos Formen lernförderlicher kognitiver Verarbeitung ausgelöst werden. Insbesondere die Forschung zu Kooperationsskripts legt nahe, dass durch eine geeignete Strukturierung der Interaktion ein solcher Austausch zwischen Schülerinnen und Schülern angestoßen werden kann (Kollar et al. 2006, Fischer et al. 2013, Vogel et al. 2017). So könnten die Schülerinnen und Schüler etwa dazu angeregt werden, in Abhängigkeit vom bereits erzielten und kooperativ diagnostizierten Verständnis der Lerninhalte gemeinsam mit dem Video zu interagieren (z. B. durch Zurückspulen, Pausieren, Reduzieren der Wiedergabegeschwindigkeit) und dadurch Fachwissensinhalte zu reorganisieren und mit ihrem Vorwissen zu verknüpfen (Renkl, 2008). Durch diese Form der Unterstützung werden möglicherweise zugleich kooperative Kompetenzen der Rezeption von Erklärvideos aufgebaut.

Anknüpfend an diese Überlegungen wird daher folgende Fragestellung untersucht: Welchen Effekt hat das durch strukturierende Hinweise unterstützte kooperative Rezipieren von Erklärvideos auf den Erwerb von Fachwissen und kooperativen Kompetenzen? Es wird angenommen, dass die Schülerinnen und Schüler, die Erklärvideos auf diese Weise kooperativ rezipieren, sowohl mehr Fachwissen als auch ein höheres Maß an kooperativen Kompetenzen erwerben. Daran anknüpfend wird explorativ der Fragestellung nachgegangen, welche kooperativen Lernaktivitäten bei der gemeinsamen Rezeption von Erklärvideos sich als besonders lernförderlich erweisen.

Dazu wird eine quasiexperimentelle Interventionsstudie mit einem Pre-Post-Kontrollgruppendesign in zwei Klassen einer niedersächsischen Oberschule durchgeführt. Die Unterrichtseinheit *A trip to London* soll die Schülerinnen und Schüler befähigen, über eine fiktive zukünftige Reise nach London zu sprechen, indem sie die Zeitform *will future* verwenden und verschiedene Sehenswürdigkeiten benennen, und beinhaltet je ein Erklärvideo zu den Themen *will future* und Sehenswürdigkeiten in London. In der Kontrollbedingung bearbeiten die Schülerinnen und Schüler die Erklärvideos individuell. In der Experimentalbedingung bearbeiten sie diese in Zweiergruppen; ihre Interaktion wird dabei durch entsprechende Hinweise auf Arbeitsblättern strukturiert. Fachwissen und kooperative Kompetenzen werden im Pre- und Posttest mit Single-Choice- und offenen Fragen erfasst. Der Einfluss der Lernsettings auf den erreichten Wissens- und Kompetenzstand wird durch Kovarianzanalysen analysiert. Darüber hinaus werden Audioaufnahmen der Interaktionen bei der kooperativen Rezeption der Erklärvideos in der Experimentalbedingung explorativ hinsichtlich der kooperativen Lernaktivitäten ausgewertet.

Die Ergebnisse versprechen Erkenntnisse über Möglichkeiten zur Förderung der Rezeption von Erklärvideos, insbesondere zu produktiven kooperativen Lernaktivitäten, an denen dabei angesetzt werden könnte. Die Identifikation von Merkmalen lernförderlicher Rezeption könnte so die Gestaltung zukünftiger Kooperationsarrangements verbessern.

63. Implementation digitaler Übungsaufgaben in die Übung der Organischen Chemie

Katrin Schüßler, Michael Giese, Maik Walpuski
Universität Duisburg-Essen, Germany

Theoretischer Hintergrund

Bisher fehlten Möglichkeiten, um in der organischen Chemie typische Aufgabenformate digital umzusetzen. Digitale Lernumgebungen können zusätzliche Lern- und Feedbackgelegenheiten schaffen und speziell vorwissensschwache Studierende, die in der regulären Lehrveranstaltung ihren Wissensrückstand nicht aufholen können (Averbeck, 2020) und gefährdet sind, die Veranstaltung nicht (in Regelstudienzeit) zu bestehen (Halikari & Nevgi, 2010), unterstützen. Für die Konzeption einer lernförderlichen Lernumgebung ist zu beachten, dass Arbeitsgedächtniskapazität begrenzt ist und diese Kapazität für den Lernprozess und nicht für unnötige Verarbeitungsprozesse zur Verfügung stehen sollte (z. B. Mayer, 2009; Pass & Sweller, 2014; Sweller, 2005). Ziel ist daher die Entwicklung einer digitalen Lernumgebung, die (1) möglichst wenig unnötige kognitive Belastung durch die digitale Umsetzung hervorruft (Mayer, 2009; Paas & Sweller, 2014) und (2) einen didaktischen Mehrwert (z. B. durch automatisiertes fehlerspezifisches Feedback; Johnson & Priest, 2014) gegenüber der aktuellen paper-pencil Lösung bietet. Hierfür wurde in das E-Assessment Tool JACK (z. B. Striewe, 2016) der Moleküleditor KEKULE.JS (Jiang, Jin, Dong, & Chen, 2016) integriert, so dass es möglich ist, Molekül-Zeichenaufgaben digital umzusetzen, automatisiert auszuwerten und Feedback zu geben.

In einer ersten Erprobung des Tools haben Studierende, die im Semester mit papierbasierten Aufgaben gelernt hatten, 8 Zeichenaufgabentandems mit jeweils einer papierbasierten (Zeichnen Sie ein 3-Ethyloctan-Molekül) und einer analog konstruierten, digitalen Aufgabe (Zeichnen Sie ein 2-Methylheptan-Molekül) bearbeitet. Für die Aufgaben wurden Personenfähigkeit sowie kognitive Belastung (Leppink, Paas, Van der Vleuten, & van Merriënboer, 2013) verglichen. Zusätzlich wurde die Usability (Brooke, 2014; Hauck, Melle & Steffen, 2022) der digitalen Aufgaben betrachtet. Der Vergleich der Aufgaben zeigt, dass für Studierende, die im Rahmen ihrer Lehrveranstaltung lediglich papierbasierte Übungsaufgaben bearbeitet habe, die digitalen Zeichenaufgaben signifikant schwieriger sind und dass das digitale Format eine signifikant höhere unnötige kognitive Belastung verursacht (Autoren, in Vorb.). Darüber hinaus stimmen Studierende, die den Aussagen zur aufgabenbedingten kognitiven Belastung stärker zustimmen, den Aussagen zur Usability seltener zu (und anders herum). Der negative Zusammenhang zwischen aufgabenbedingter kognitiver Belastung und Usability zeigt, dass das Studierende, die bereits durch die Aufgaben herausgefordert sind, die Usability des (unvertrauten) digitalen Formats als geringer empfinden.

Fragestellung

Im Wintersemester wird das digitale Tool nach einer ausführlichen Einführung systematisch in die Übung der Organischen Chemie integriert, so dass Studierende bereits in der Lernphase mit dem Tool arbeiten und nicht mehr ausschließlich auf das Papierformat trainiert werden.

(1) Genügt eine längere Übungsphase, um Unterschiede in der Personenfähigkeit und der unnötigen kognitiven Belastung zwischen den Formaten zu verringern oder verursacht das digitale Tool grundsätzlich eine höhere unnötige kognitive Belastung?

(2) Bleibt der negative Zusammenhang zwischen aufgabenbedingter kognitiver Belastung und Usability stabil oder verschwindet er nach einer Gewöhnungsphase?

Zusätzlich ist für den Einsatz des Tools in der Übung ein Ausbau der Feedbackfunktion geplant, um Studierende in ihrem Lernprozess zu unterstützen (Johnson & Priest, 2014).

Methode

Die Organische Chemie-Übung wird im Wintersemester von Chemie-Lehramtsstudierenden im dritten Semester besucht (*N* = 30-40). Für die Untersuchung werden erneut Aufgabentandems verwendet, die aus dem regulären Übungsaufgabenpool heraus entwickelt, im Rahmen der Übung bearbeitet und IRT-skaliert (eindimensionales Rasch-Modell, WINSTEPS) werden. Kognitive Belastung (Leppink, Paas, Van der Vleuten, & van Merriënboer, 2013) und Usability (Brooke, 2014; Hauck, Melle & Steffen, 2022) werden wieder über 9-stufige Ratingskalen erhoben.

Ergebnisse und ihre Bedeutung

Die ersten Ergebnisse haben gezeigt, dass neue Formate neue Herausforderungen mit sich bringen und vorsichtig eingeführt werden müssen, da sie speziell schwächere Studierende zusätzlich herausfordern können. Gleichzeitig ist das neue Tool durch zusätzliche Übungs- und Feedbackgelegenheiten speziell für schwächere Studierende vielversprechend. Die forschungsbegleitete Implementation in die Übung verspricht evidenzbasierte Einblicke in das (veränderte) Lernen der Studierenden mit dem Tool.

64. Digital natives = digital experts? Designing a novel instrument to measure digital competence in NEPS-SC8 secondary school students

Sümeyra Tural, **Mariann Schwaß**
Leibniz-Institut für Bildungsverläufe, Germany

With the digitization of modern day society, the competence of handling information and communication technologies (ICT) has become increasingly significant. Newer generations that grow up in a continuously more digitized world from a young age and almost inevitably familiarize themselves with the technologies and devices of this new environment, are labelled as ‘digital natives’ (Prensky, 2001), ‘generation internet’ or ‘whiz-kids’ (Lee & Chae, 2007). Often, these terms are used to imply a certain level of competence in operating with the internet and online media that is thought to be automatically acquired by frequent usage. On the other hand, many pieces of anecdotal and scientific evidence indicate the challenge that the manifold information of modern media and the internet can pose for growing children and adolescents. “They are not only called ‘whiz-kids’ […] but also ‘risk-kids’”, state Valcke et al. (2010) and emphasize the importance of teaching children a critical and reflected approach.

In order to capture the aspect of critical reflection and social communication within the internet, a new test to measure ‘digital competence’ has been developed for Starting Cohort 8 of the NEPS. The test aims to reflect the knowledge and abilities that are needed to responsibly approach the possibilities but also the risks and consequences of the ubiquitous accessibility of the internet and to reflect not only on the *content* of digital media, but also to identify their modes of action. In this way, the newly developed test complements the aspect of handling technology and programmes that is already established in the NEPS and that is represented by the competence domain ‘ICT’.

This poster presentation aims to introduce the conceptual framework for this novel test domain. In a synthesis and extension of former research (e.g. Calvani, Fini & Ranieri, 2009; Carreto, Vuorikari & Punie, 2017; Digital Dannelse, 2009; Fraillon et al., 2019), four main facets with multiple subfacets were defined for the digital competence test of the NEPS: (1) *Information competence,* (2) *Detection of intentions and strategies*, (3) *Communication and interaction competence,* and lastly, (4) *Data handling competence.*

The computer based test to measure digital competence will first be applied in a pilot study for the NEPS-SC8 in 2022. Data will be gathered in regular schools and schools for children with learning difficulties, in grade 6 and 8, respectively. In the long run, the NEPS data will allow to investigate the development, influencing factors and consequences of digital competence in secondary school children.

65. Gelingensbedingungen für die Nutzung digitaler Medien von Lehramtsstudierenden in der Schulpraxis: Ein Critical Review

Marcel Capparozza[1,2], Jessica Kathmann[1], Gabriele Irle[1]
[1]Leibniz-Institut für Wissensmedien (IWM); [2]Institut für Bildungsanalysen Baden-Württemberg (IBBW)

Eine zentrale Aufgabe des Lehramtsstudiums ist die Qualifizierung zukünftiger Lehrkräfte für die Nutzung digitaler Medien im Unterricht (Monitor Lehrerbildung, 2021). Um mit digitalen Medien unterrichten zu können, sollten Lehramtsstudierende lernen, wie sie technisches, pädagogisches und fachwissenschaftliches Wissen miteinander verknüpfen (*TPACK* bzw. *Technological Pedagogical Content Knowledge;* Mishra & Koehler, 2006). Praktika ermöglichen Lehramtsstudierenden durch Beobachtungs- und Reflexionsphasen, Unterrichtspraxis und der Teilnahme an schulischen Aktivitäten das Erleben erster authentischer Erfahrungen im Klassenzimmer (Cohen, Hoz & Kaplan, 2013). Authentische Erfahrungen stellen eine zentrale Strategie für die Vorbereitung auf das Unterrichten mit digitalen Medien dar (Kay, 2006; Røkenes & Krumsvik, 2014; Tondeur et al., 2012). In diesem Beitrag wird mithilfe eines Critical Review thematisiert, wie individuelle Voraussetzungen (Fragestellung 1), Rahmenbedingungen an Hochschulen (Fragestellung 2) sowie Merkmale der Praktikumsgestaltung bzw. schulische Rahmenbedingungen (Fragestellung 3) den Einsatz digitaler Medien in der Schulpraxis fördern bzw. erschweren. Außerdem werden motivationale und kognitive Wirkungen von schulpraktischen Unterrichtserfahrungen mit digitalen Medien vorgestellt (Fragestellung 4).

Die Grundlage für dieses Critical Review bildet eine deutsch- und englischsprachige Literaturrecherche in den Datenbanken Fachportal Pädagogik, ERIC, Web of Science, ERC sowie LearnTechLib im Suchzeitraum 01/2016 bis 08/2021 mit 2844 Suchergebnissen. Für diese Suche wurden Suchbegriffe zum Lehramtsstudium (z. B. *teacher education*) mit Suchbegriffen zu Digitalisierung (z. B. *technology*) und zur Schulpraxis (z. B. *field experience*) kombiniert. Ein mehrstufiger Such- und Auswahlprozess (Prüfung der Titel und Abstracts, Prüfung der Volltexte, Rückwärts- und Vorwärtssuche) führte zur Auswahl von 38 Studien aus peer-reviewten Zeitschriften.

Die Auswertung der Studienergebnisse zeigt, dass drei Gelingensbedingungen Lehramtsstudierende beim Unterrichten mit digitalen Medien in der Schulpraxis unterstützen:

(1) Individuelle Voraussetzungen der Lehramtsstudierenden: Dazu gehören u. a. positive Einstellungen zum Unterrichten mit digitalen Medien (z. B. Goldstein & Tessler, 2017; Liu, 2016; Sadaf, Newby & Ertmer, 2016), eine konstruktivistische Haltung (Han, Shin & Ko, 2017) sowie Kompetenzen und Vorerfahrungen (z. B. Aslan & Zhu, 2017; Liu, 2016; Sun, Strobel & Newby, 2017).

(2) Rahmenbedingungen an Hochschulen: Studien zeigen beispielsweise, dass im Studium erworbene Kompetenzen mit der Häufigkeit der Mediennutzung zusammenhängen (Aslan & Zhu, 2017; Goldstein & Tessler, 2017; Liu, 2016). Gefördert werden können diese durch Hochschulpraktiken wie z. B. Microteaching (Ünal et al., 2017) und ausreichend lange Explorationsphasen (Reese et al., 2016).

(3) Merkmale der Praktikumsgestaltung und schulische Rahmenbedingungen: Lehramtsstudierende benötigen ausreichend Zeit für das Unterrichten mit digitalen Medien (z. B. Cullen & Karademir, 2018; Jones et al., 2017; Laherto & Laherto, 2018) und Klassenzimmer, die mit funktionsfähiger Soft- und Hardware und Infrastruktur ausgestattet sind (z. B. Batane & Ngwako, 2017; Krause & Lynch, 2018; Trainin et al., 2018). Außerdem profitieren sie von Mentor:innen, die sie bei der Nutzung digitaler Medien unterstützen oder ihnen ausreichend Autonomie für die Gestaltung des Unterrichts einräumen (Cullen & Karademir, 2018; Jones et al., 2017). Ein Schlüsselfaktor scheint zudem die mediendidaktische Kompetenz der betreuenden Lehrkräfte zu sein (Nelson, 2017; Nelson & Hawk, 2020).

Die Analyse der Studien zeigt außerdem, dass sich die Nutzung digitaler Medien in der Schulpraxis bei den Lehramtsstudierenden sowohl mit motivationalen Faktoren wie z. B. einer höheren Selbstwirksamkeitserwartung (z. B. Han et al., 2017; Krause, 2017; Reese et al., 2016) als auch mit einem Zuwachs an Kompetenzen (z. B. Agyei & Voogt, 2016; Buss et al., 2018; Sun et al., 2017) assoziiert ist.

Aus den Ergebnissen lassen sich evidenzbasierte Empfehlungen zur Gestaltung von Praktika in der Ausbildung von Lehrkräften ableiten: (1) Stärkung der Ausbildung an Hochschulen mit Blick auf mediendidaktische Kompetenzen und Einstellungen, (2) Erfassung und Berücksichtigung individueller Voraussetzungen während der Schulpraktika, (3) Ausstattung von Universitäten und Praktikumsschulen mit moderner und funktionsfähiger Soft- und Hardware und (4) Stärkung der Vorbildfunktion von Lehrenden und Mentor:innen in Bezug auf die Nutzung digitaler Medien (z. B. durch Fortbildungen).

66. Unterstützung von Experimentierprozessen in der Grundschule durch digitale Medien

Saskia Knoth, Michael Haider, Astrid Rank
Universität Regensburg, Deutschland

Theoretischer Hintergrund

Im naturwissenschaftlichen Sachunterricht treten an verschiedenen Stellen des Experimentierprozesses Schwierigkeiten auf (u.a. Hammann et al., 2006; Grube, 2010; Baur, 2018). Es bedarf vielfältiger Kompetenzen und Arbeitsweisen, um den Prozess der Erkenntnisgewinnung erfolgreich zu durchlaufen (Haider & Fölling-Albers, 2013; Haider & Munser-Kiefer, 2019). Der Einsatz digitaler Medien eröffnet hier die Möglichkeit der adaptiven Unterstützung und kann gleichzeitig der

Binnendifferenzierung dienen (Eickelmann, 2010). Die Vorhersage des Lernerfolgs beim Einsatz digitaler Medien ist stark vom thematischen, aber auch medienbezogenen Vorwissen abhängt (Herzig, 2014). Selbststeuerung und der Einsatz von Lernstrategien nehmen außerdem Einfluss auf den Lernerfolg (ebd.). Bezüglich der Gestaltung von Hilfsangeboten wirkt sich die Steuerung und Kontrolle durch den Lernenden positiv auf das Lernergebnis aus (Karich et al., 2014). Lernumgebungen mit vielen Wahloptionen werfen bei Lernenden mit geringem Vorwissen Probleme auf (Tergan, 2002). Besonders wirksam ist der Einsatz gestufter Hilfen im Sinne der „guided inquiry“, wobei mündliche Hilfestellungen weniger kognitive Anstrengung (Extraneous Cognitive Load) verursachen als schriftliche (Schmidt et al., 2019). Die Umsetzung der Unterstützung wird im Projekt auf den verschiedenen Stufen des praxisorientierten SAMR-Modells (Puentedura, 2006) variiert. Der Modelltitel steht für die einzelnen Stufen, wobei von der untersten Stufe Substitution über Augmentation und Modification bis zur Redefinition eine Steigerung des Einsatzes digitaler Möglichkeiten erfolgt. Puentedura verbindet damit, dass sowohl die Qualität wie auch der pädagogische Nutzen mit dem Stufenanstieg zunehmen (Puentedura, 2006).

Messmethodisch gibt es zwei gängige Verfahren für die Erhebung von experimenteller Kompetenz: Paper & Pencil-Tests oder praktische Tests. Verschiedene Autoren (u.a. Hammann, 2004) zeigen auf, dass experimentelle Kompetenz immer in konkreten inhaltlichen Situationen gezeigt wird und nicht losgelöst von den naturwissenschaftlichen Inhalten erhoben werden kann. Der Kompetenzgewinn soll hier in Anlehnung an die Tests des Konsortium HarmoS Naturwissenschaften+ (2010) gemessen werden. Als Forschungsdesiderat ergibt sich die Frage, inwieweit die Defizite in der experimentellen Kompetenz bei Grundschülerinnen und Grundschülern über adaptive punktuelle digital gestützte Unterstützungsmaßnahmen im Experimentierprozess aufgegriffen werden können und experimentelle Kompetenz gefördert werden kann.

Fragestellung

Im Fokus des Posters steht die Frage, ob Schülerinnen und Schüler durch digitale Medien beim Experimentieren unterstützt werden können, sodass sie hinsichtlich des Lernzuwachses, der experimentellen Kompetenz und der Motivation profitieren können. Dementsprechend werden u.a. folgende Fragestellungen und Hypothesen als Erkenntnisziele des Projektes formuliert:

1. Lassen sich Experimente bzgl. des Lernzuwachses, der Motivation und der experimentellen Kompetenz durch digitale Medien unterstützen? --> Hyp. 1: *Experimente lassen sich generell durch Unterstützungsmaßnahmen digitaler Medien stärken, sodass ein Zuwachs des Wissens, der Motivation und der experimentellen Kompetenz deutlich wird.*
1. Haben Lernunterstützungen auf verschiedenen Ebenen des SAMR-Modells unterschiedliche Effekte auf Lernzuwachs, Motivation und experimentelle Kompetenz? --> Hyp. 2: *Die Lernunterstützung auf den Ebenen M und R des SAMR-Modells hat einen größeren Effekt auf Lernzuwachs, Motivation und experimentelle Kompetenz aufgrund der besser umsetzbaren adaptiven Hilfestellungen.*

Methode

Die quantitative Interventionsstudie mit quasiexperimentellem Cross-Over-Design mit Kontrollgruppe in 12 Klassen des dritten Schuljahres (N=250) erhebt Lernzuwachs, Motivation und experimentelle Kompetenz sowie deren Veränderung durch einen digital unterstützten Unterricht. Die Umsetzung der Unterstützung variiert dabei auf den Stufen des SAMR-Modells. Geprüft wird u.a. der Einfluss dieser Variation auf die unabhängigen Variablen. Neben den motivationalen sowie kognitiven Auswirkungen muss auch die Frage nach der kognitiven Belastung der Kinder mit unterschiedlichen

Lernvoraussetzungen fokussiert werden. Ausgewählte varianz- und regressionsanalytische Ergebnisse werden präsentiert und diskutiert.

Ergebnisse & ihre Bedeutung

Alle vier Unterrichtsstunden bewirken einen signifikanten Lernzuwachs mit unterschiedlich stark ausgeprägter Effektstärke. Die Veränderung der Motivation zeigt Tendenzen, jedoch keinen signifikanten Zeiteffekt. Die digitale Unterstützung scheint keinen Einfluss auf die experimentelle Kompetenz zu haben. Die Ergebnisse weisen darauf hin, dass der Unterrichtsinhalt entscheidender ist. Unterschiedliche Auswirkungen der medialen Unterstützung werden dargestellt.

67. Vertreten kann doch jede*r! – Herausforderungen studentischer Vertretungslehrkräfte im Umgang mit einer heterogenen Schülerschaft

Sarah Rau-Patschke, Susanne Gerlach
Universität Duisburg-Essen, Germany

Heterogenität erleben alle Lehrkräfte in unterschiedlichem Ausmaß. Dem kann im Unterricht divergent begegnet werden. Neben verschiedener, auf Förderschwerpunkte spezifisch ausgerichteter Maßnahmen, gibt es auch Konzeptualisierungen, die dem hier fokussierten, weiten Inklusionsbegriff entsprechend Hilfestellungen bereithalten, z.B. das Universal Design for Learning (CAST, 2011).

Der Umgang mit Heterogenität (UmH) ist in allen Phasen der Lehrkräftebildung als herausfordernd zu bezeichnen: Die Arbeitsgruppe um Baer zeigte, dass der UmH in frühen Praxisphasen nur wenig Berücksichtigung findet (2011). Für den Vorbereitungsdienst bestätigen Untersuchungen (Windt et al., 2017; Rau, 2017), was Pietsch bereits 2010 für ausgebildete Lehrkräfte nachzeichnete: Der UmH als Merkmal qualitätsvollen Unterrichts (Berliner, 2005) kommt überhaupt erst dann von Lehrkräften im Unterricht zum Tragen, wenn Aspekte der Unterrichtsorganisation und -strukturierung in ausreichendem Maß umgesetzt werden (Pietsch, 2010). Im Gegensatz zur Basisdimension Klassenführung (Klieme et al. 2006) zeigt sich, dass der UmH in geringen Umfang im (Sach-)Unterricht umgesetzt wird (Windt et al., 2017). Auch zum Ende des Vorbereitungsdiensts bleiben diese Unterschiede bestehen (Rau, 2017). Dass der UmH insbesondere in Deutschland ein noch eher wenig präsentes Thema in der Lehrkräfteausbildung der naturwissenschaftlichen Fachdidaktiken ist, könnte ein Grund für die unzureichende Professionalisierung in diesem Gebiet sein (Markic & Abels, 2014).

Darüber hinaus herrscht ein flächendeckender Lehrkräftemangel (DESTATIS, 2021). Dieser manifestiert sich insbesondere im naturwissenschaftlichen Unterricht. Dieser Personalmangel wird versucht, über verschiedene Wege zu kompensieren. Neben Seiten- bzw. Quereinsteiger*innen (Anders, 2022) werden zunehmend studentische Vertretungslehrkräfte (sVLK) an Schulen eingesetzt (Lobert & Pfitzner, 2021; Rau-Patschke, 2022). Das Thema der sVLK wird erst von wenigen Studien in den Blick gerückt (ebd.). Als problematisch bezeichnet dabei Freiling, dass insbesondere der Unterrichtsausfall an Schulen mit hoher Heterogenitätsspanne durch Vertretungslehrkräfte abgedeckt wird (2020). Gerade Schüler*innen mit schwierigen Ausgangsbedingungen profitieren aber von gut ausgebildeten Lehrkräften, sodass hier mehrere bildungsbezogene Herausforderungen aufeinandertreffen. Das wiederum führt unter Umständen zu einem hohem Belastungserleben der

sVLK und kann in Überforderung oder Zweifel an der Studiengangswahl münden (Rau-Patschke, 2022).

Am Beispiel des naturwissenschaftlichen Unterrichts wird untersucht, wie die sVLK ihr Verständnis hinsichtlich der Förderschwerpunkte und Heterogenitätsdimensionen einschätzen, inwiefern sie mit Heterogenität konfrontiert sind, welche Herausforderungen sie erleben und welche Unterstützungsbedarfe sich daraus ergeben.

Dem Forschungsinteresse wird mit einer Online-Fragebogenerhebung nachgegangen. Die sVLK werden zu ihrer Arbeits- und Studiensituation (1), zu Aspekten des Unterrichtens (2) sowie ihre Zufriedenheit, Selbstwirksamkeitserwartung hinsichtlich ihrer Tätigkeit allgemein (3) und im Besonderen im UmH (4) befragt. Die Fragen der Blöcke (1)-(3) lehnen sich an erprobte Instrumente an. Der Fragenblock zum UmH (4) ist selbst konstruiert und weißt akzeptable (α=.68) bis sehr gute interne Konsistenzen (α=.92) auf. Insbesondere in Ergänzung zu Frageblock (4) vertiefen fünf leitfadengestützte Einzelinterviews die Einsicht in die Arbeit mit heterogenen Lerngruppen durch sVLK.

Die gewonnenen Daten der Fragebogenerhebung werden in SPSS transferiert und aufbereitet. Der Fragenbogen befindet sich derzeit im Einsatz (bislang n=23). Die Interviews werden kategorisiert (induktive Kategorienbildung), zu 10% doppelt codiert (Lamnek & Krell, 2010) und skalierend-strukturiert (Mayring, 2015) in SPSS eingepflegt. Die Auswertung erfolgt deskriptiv, explorativ. So können im bislang wenig beforschten Feld erste Zusammenhänge entdeckt und Hypothesen generiert werden (Tukey, 1977).

Erste Einblicke in die Daten zeigen, dass 91,3% der sVLK die bezahlte Praxiserfahrung und 65,2% den Transfer von Theoriewissen in die Schulpraxis schätzen. Die sVLK beurteilen ihr Vorwissen zu Heterogenitätsdimensionen eher gut (MW=1,98; SD=.49; Skala 0-3) und geben an, dass die Realisation im Unterricht eher gut gelingt (MW=2.00; SD=.48) und sie sehr häufig damit in Kontakt treten (MW=2,38; SD=.43). Die Einschätzungen hinsichtlich der Förderschwerpunkte fallen in allen drei Bereichen signifikant schlechter aus (p<.001). Hinsichtlich der Lehramtstypen lassen sich diesbezüglich keine Gruppenunterschiede feststellen, sodass davon auszugehen ist, dass der UmH sowie den Förderschwerpunkten ein Thema für alle sVLK ist.

68. Wirkungen studienunabhängiger Unterrichtstätigkeit auf Studierende des Lehramts an Grundschulen (PRAX-Studie)

Daniel Rehfeldt, Hilde Köster
Freie Universität Berlin, Germany

Die studienunabhängige Unterrichtstätigkeit – das Arbeiten als (Vertretungs-)Lehrkraft neben dem Lehramtsstudium – liegt im Trend (z. B. Ibrahim & Baumann, 2021). Dabei überrascht sowohl die hohe Quote dieser Studierenden von oft mehr als 30% (z. B. Bröder et al., 2022) als auch die Anzahl der von ihnen zu leistenden Wochenstunden – und dies mitunter bereits im ersten Studienjahr (> 12h, Bäuerlein, 2018).

Vor dem theoretischen Hintergrund des Angebots-Nutzungs-Modells zu den Lerngelegenheiten im Lehramtsstudium und in außeruniversitären Bereichen (Doll et al., 2020) ergibt sich vor dem großen

Umfang der „Nutzung“ dieses sehr intensiven außeruniversitären Lern-„Angebots“ (Hascher, 2007) die Frage nach dessen Wirkung auf die Studierenden.

Die wenigen deutschsprachigen Forschungsergebnisse zu studienunabhängiger Unterrichtstätigkeit stammen aus dem Schweizer Lehrkräftebildungssystem: Studierende bewerten die Erfahrungen als positiv, allerdings mit höherer Intensität als Praktika (Bäuerlein, 2018; Hascher, 2007). Zentrale Herausforderungen sind die Anforderungen einer „doppelten Professionalisierung“ (Helsper, 2001) bei hoher Intensität und Selbstorganisationsaufwand. Hinzu kommen Probleme beim Klassenmanagement durch den artifiziellen Status und die Unerfahrenheit (Scheidig & Holmeier, 2022; Hascher, 2007). Bei professionellen Kompetenzen ergaben sich negative Zusammenhänge zur eigenen Kompetenzeinschätzung (Scheidig & Holmeier, 2022) und fehlende Entwicklungen im pädagogischen Wissen oder Interesse (Depping et al., 2021). Entgegen den Erwartungen fanden Scheidig und Holmeier (2022) nur kleine Einflüsse auf eine stärkere Praxis-Erwartung im Studium, allerdings mit methodischen Einschränkungen durch die Nutzung von Einzelitems.

Als förderlich für professionelle Kompetenzen auf Angebotsseite wurden (analog zu Praktika) Betreuung bzw. Kollegium (Scheidig & Holmeier, 2022; Bäuerlein, 2018; Hascher, 2007) und Reflexion herausgestellt. Die Herausforderungen bei ausbleibender Betreuung (betrifft 82% bei Bäuerlein, 2018) erscheinen immens, Studierende berichteten von z. B. großen Problemen beim Klassenmanagement und wenig Feedback. Die Attribuierung der Probleme geschah zudem oft auf die Lernenden gerichtet (Hascher, 2007) – dies zeichnet ein Bild geringer professioneller Entwicklung. Reflexionen scheinen zudem seltener als in Praktika und insgesamt wenig relevant zu sein (Hascher, 2007; Doll et al., 2020). Doll et al. (2020) betonen hier den Noviz:innenstatus der Studierenden (Berliner, 2004) und die hohe Wichtigkeit einer adäquaten Betreuung.

Es ergeben sich daher zusammenfassend folgende Fragestellungen: Es ist unklar, welche Rolle die studienunabhängige Unterrichtstätigkeit für (1) das Nutzungsverhalten der Studierenden (Studium, Unterricht) und (2) die Überzeugungen der Studierenden spielt. Relevant scheint dafür die (3) Qualität und Intensität von Betreuung und Theorie-Praxis-Reflexion. Insbesondere für das deutsche Bildungssystem fehlen bislang Erkenntnisse dieser Art.

Das Design der Studien des PRAX-Projekts zur Schließung dieser Forschungslücken umfasst in der beitragsrelevanten Studie 1 einen Quasi-Längsschnitt und in Studie 2 einen echten Längsschnitt.

Methodisch wird in Studie 1 im zweiten und achten Fachsemester des Studiengangs Lehramt für die Grundschule im Sommersemester 2023 per Rating-Fragebogen eine Vollerhebung im Sachunterricht angestrebt (N ≈ 700).

Erhoben werden auf Angebotsseite schulseitig die Qualität und Quantität der studienunabhängigen Unterrichtstätigkeit inklusive Betreuung (Bäuerlein, 2018), universitätsseitig die Qualität der theoretischen Studieninhalte für die Praxis (Scheidig & Holmeier, 2022; Dübbers, 2022) und die Nutzung der Lerngelegenheiten im Studium bzw. in der Schule (Mertens et al., 2020). Zur Erforschung der Überzeugungen werden die Offenheit gegenüber Theorie (Osterheert, 2001), transmissive Überzeugungen (Kunina-Habenicht et al., 2013) und Offenheit der Lernsteuerung (Hesse & Latzko, 2009) untersucht und – für einen Teil der Stichprobe – mit einem impliziten

Assoziationentest (Mierke, 2005) bzgl. offener versus geschlossener Lehrformen flankiert. Für die Reflexionskompetenz werden die Skalen nach König et al. (2014) verwendet und mit Informationen zur Häufigkeit und Gegenstand reflexiver Tätigkeiten kombiniert.

Die Auswertung erfolgt u.a. multivariat als Quasi-Längsschnitt. Hypothetisiert werden kann eine höhere Theorieskepsis und Valorisierung praxisnaher universitärer Lernangebote bei Studierenden mit zunehmender Dauer der studienunabhängigen Unterrichtstätigkeit.

Im Poster wird das genaue Studiendesign vorgestellt und diskutiert.

69. Einstellungen zur emotionalen Unterstützung von Studierenden im Praxissemester: (Auch) eine Frage der Wahrnehmung von MentorInnen?

Nicole Bosse[1], Florentine Hickethier[1], Alexander Gröschner[1], Bernadette Gold[2]
[1]Friedrich-Schiller-Universität Jena, Germany; [2]Technische Universität Dortmund

Theoretischer Hintergrund

Studierende nehmen MentorInnen als wichtige Einflusspersonen im Praktikum wahr (Izadinia, 2015). MentorInnen werden dabei oftmals als Rollen-Modelle und Vorbilder unterrichtlichen Handelns erachtet (u.a. Gröschner & Hascher, 2022; Kang, 2021). Vor dem Hintergrund der Bedeutung einer positiven Lehrer-Schüler-Beziehung (Brekelmans et al., 2002) wurden die Daten einer Befragung von 310 Praxissemesterstudierenden der Grundschule und weiterführender Schulen daraufhin untersucht, welche Einstellungen die Studierenden in Bezug auf den Aspekt der emotionalen Unterstützung von SchülerInnen besitzen und inwieweit sich die Einstellungen im Verlauf des Praxissemesters durch den Zugewinn an Erfahrungen verändern. Darüber hinaus wurde untersucht, inwiefern dieser Aspekt bedeutsam ist, wenn die Studierenden die MentorInnen-Perspektive als Gegenstand der Unterstützung durch die MentorInnen wahrnehmen.

Folgende Fragestellungen stehen im Fokus der Studie:

- Verändern sich die Einstellungen von Praxissemesterstudierenden in Bezug auf die emotionale Unterstützung im Verlauf des Praktikums?
- Welchen Einfluss nimmt der/die MentorIn auf die Einstellungen?
- Inwieweit zeigen sich Unterschiede zwischen Studierenden des Grundschullehramts und der Sekundarstufe?

Methodisches Vorgehen

Die Daten wurden im Rahmen einer Befragung zum mittleren Zeitpunkt (t1) und am Ende des Praxissemesters (t2) im SoSe 2021 und WiSe 2021/22 erhoben. Bei den StudienteilnehmerInnen handelte es sich um N = 310 Studierende (weiblich = 237 Studierende, männlich = 67 Studierende; Alter: M = 23.72, SD = 2.81) zweier Universitäten.

Einerseits wurden am letzten Messzeitpunkt die studentischen Einstellungen zur emotionalen Unterstützung (Gold & Müller, 2021, in Vorbereitung; Bowlby, 1969; Cornelius-White, 2008; Pianta et al., 2012), andererseits die studentische Einschätzung der Bedeutung dieser Aspekte für den/die MentorIn erhoben (Gold & Müller, 2021, in Vorbereitung). Die studentische Einstellung zur emotionalen Unterstützung von SchülerInnen wurde mit acht Items erfasst (Bsp.: „In meinem Unterricht achte ich besonders darauf, dafür zu sorgen, dass meine SchülerInnen sich im Unterricht wohl fühlen.“; 6-stufige Likert-Skala von 1 = „stimme überhaupt nicht zu“ bis 6 = „stimme voll und ganz zu“). Die interne Konsistenz der Skala zum Prä- und Posttest lag in einem guten Bereich (α_{t1} = .87 /

α_{t2} = .88). Die von den Studierenden wahrgenommene Beziehung des/der MentorIn zu den SchülerInnen wurde mit sieben Items erhoben, die wiederum auf einer 6-stufigen Likert-Skala einzuschätzen waren (Bsp.: „In ihrem Unterricht hat meine Betreuungslehrkraft die Bedürfnisse der SchülerInnen berücksichtigt.“; 1 = „stimme überhaupt nicht zu“ bis 6 = „stimme voll und ganz zu“; α = .90).

Ergebnisse

Die Studierendeneinstellungen verändern sich tendenziell positiv zum Ende des Praxissemesters ($t(309) = 3.84$, $p < .001$; $d = 0.22$). Nach Cohen (1988) liegt hier ein kleiner Effekt vor.

Zum letzten Messzeitpunkt kann ein moderat positiver Zusammenhang zwischen der studentischen Einstellung zur emotionalen Unterstützung und der wahrgenommenen Bedeutung des Aspekts für den/die MentorIn festgestellt werden ($r = .40$, $p < .001$). Für die Primar- und Sekundarstufe zeigen sich dabei keine bedeutsamen Unterschiede für den Zusammenhang dieser beiden Konstrukte (Primarstufe: $r = .37$, $p < .001$; Sekundarstufe: $r = .31$, $p < .01$).

Die gewonnenen Ergebnisse spiegeln die subjektiven Überzeugungen und Wahrnehmungen der Studierenden wider. Die Ergebnisse können erste Hinweise für den Einfluss der Wahrnehmung der Betreuungslehrkraft als Rollen-Modell auf die eigenen Einstellungen zur emotionalen Unterstützung im Praxissemester darstellen.

Weitere Analysen im Hinblick auf die Veränderungen der Einstellungen im Verlauf des Praxissemesters sowie mögliche Einflussfaktoren sind geplant.

70. Bedeutung von Theorie-Praxis-Verknüpfungen im Feld der Lehrkräftebildung: Eine Metaanalyse

Anna-Lena Gérard[1], Larissa Legler[1], Alexander Kauertz[1], Dorothee Kohl-Dietrich[2], Heidrun Ludwig[1], Sandra Nitz[1], Jürgen Roth[1]

[1]Rheinland-Pfälzische Technische Universität Kaiserslautern Landau, Campus Landau; [2]Pädagogische Hochschule Karlsruhe

Eine enge Zusammenarbeit zwischen forschungsbasierter Theorie und schulischer Praxis ist im Bereich der Lehrkräftebildung elementar für gelingende Implementationen im schulischen Alltag (Fraefel & Scheidig, 2018; Rothland, 2020). So wird Theorie-Praxis-Verknüpfungen in den letzten Jahren große Bedeutung zugemessen (Hascher, 2012), beispielsweise durch vielfältige Projekte der „Qualitätsoffensive Lehrerbildung“ (Bresges et al., 2019). Wie eine solche Verknüpfung jedoch am besten gestaltet wird, bleibt durch die große Heterogenität laufender und abgeschlossener Forschung offen (Cramer, 2014). Zusätzlich wird ein umfassender Überblick erschwert, da insbesondere Ergebnisse zu gelungenen Verknüpfungen in den Fokus rücken, während nicht förderliche Zugänge selten publiziert werden. Eine Systematisierung des Forschungsstandes könnte als Grundlage dienen, die Lehrkräftebildung langfristig und phasenübergreifend zu verbessern, was folglich positive Effekte auf die Schüler*innen abbilden könnte. Hier setzt unsere Metaanalyse an. Ziel dieser ist es, die Datenlage zu Theorie-Praxis-Verknüpfungen in allen Phasen der Lehrkräftebildung systematisch zu sichten und zu analysieren sowie empirische Evidenzen für theoretische Überlegungen zusammenzustellen bzw. ggf. als Desiderat für zukünftige Forschung zu identifizieren. Wir gehen damit der Forschungsfrage nach, welche Bedeutung Theorie-Praxis-Verknüpfungen in der Entwicklung professioneller Kompetenzen von (angehenden) Lehrkräften zukommt.

Um die Datenbasis der Metaanalyse zu generieren, werden vier Datenbanken ausgewählt, welche ein breites inhaltliches Spektrum abbilden. Mithilfe einer festgelegten Anzahl sowohl deutscher als auch englischer Begriffskombinationen werden diese Datenbanken systematisch durchsucht. Weiterhin werden die Suchergebnisse anhand zusätzlicher Parameter eingeschränkt: der Zeitraum 2000 bis 2021, Zeitschriften mit Gutachterverfahren, quantitative Verfahren und eine erneute Festlegung der Sprachen. Einige Schulfächer, hier Sport, Bildende Kunst, Darstellendes Spiel und Musik, werden exkludiert, da hier ein anderer Umgang mit Theorie-Praxis-Verknüpfungen angenommen wird. Die Datenbankabfrage liefert 15495 Titel, die zur weiteren Strukturierung in ein Literaturverwaltungsprogramm importiert werden. Duplikate und Artikel, welche von den vorher festgesetzten Kriterien abweichen, werden entfernt. Die übrigen 7511 Artikel werden nach dem Kriterium des Bezugs zur Lehrkräftebildung anhand der Titel und der Kurzbeschreibungen im Vier-Augen-Konzept sortiert. Beim weiteren Vorgehen werden die 5423 Artikel nach folgenden inhaltlichen Kriterien gesichtet: a) die Regelschulzeit in Deutschland als Referenzrahmen und somit Exklusion von frühkindlicher und Erwachsenenbildung, b) Exklusion von Waldorf- und Montessori-Pädagogik im Hinblick auf unterliegende und nicht ausschließlich in universitärer Bildung erlernbare Bildungskonzepte, sowie c) Exklusion von Fallstudien. Im aktuellen Schritt werden die verbleibenden 3667 Artikel auf einen direkten Bezug zu Theorie-Praxis-Verknüpfungen in der Lehrkräftebildung überprüft.

Die bisherige Sichtung des zusammengestellten Literaturkorpus bestätigt die Annahme, dass die vorliegende Forschungslage sehr heterogen ist. Diese Bandbreite spiegelt sich beispielsweise in den Inhalten wider, welche etwa Themen wie Inklusion, aber auch Sonderpädagogik aufgreifen. Die große Bedeutung von Theorie-Praxis-Verknüpfungen zeigt sich dabei nicht auf den deutschsprachigen Raum beschränkt, sondern scheint auch in anderen Ländern vernommen zu werden, was Artikel des Korpus aus derzeit 58 Ländern aufzeigen. Während die Anzahl der Texte mit Bezug zur ersten und dritten Phase der Lehrkräftebildung ähnlich hoch ist, ist die zweite Phase nur marginal vertreten. Einerseits wird die Beziehung zwischen Theorie und Praxis in der ersten und der dritten Phase der Lehrkräftebildung unter sehr unterschiedlichen Gewichtungen betrachtet, mit einem Überhang an Theorie in der ersten und der dominierenden Praxis in der dritten Phase (Blömeke et al., 2006). Andererseits spielt die Verknüpfung dieser beiden Aspekte für beide Phasen eine entscheidende Rolle. Ergebnisse weiterer Analysen des Korpus werden im Rahmen der Postersession vorgestellt und diskutiert.

Diese Forschungsarbeit ist ein erster Schritt in die Richtung einer systematischen Sichtung des bisherigen Forschungsstandes. Sie kann hierbei als Ausgangspunkt für die Entwicklung eines theoretischen Rahmens für zielführende und kognitiv aktivierende Theorie-Praxis-Verknüpfungen gesehen werden. Zusätzlich kann sie zur Identifizierung weiterer Forschungsdesiderate dienen, beispielsweise zukünftig einen expliziten Fokus auf gescheiterte Theorie-Praxis-Verknüpfungen zu setzen. Unsere Meta-Analyse reiht sich damit in das Feld der Theorie-Praxis-Forschung ein und kann als weiterer Schritt in der Entwicklung der Lehrkräftebildung dienen.

71. Nutzung von Forschungswissen durch Lehrpersonen – Eine datengestützte Betrachtung von Nutzungstypen

Franziska Maria Locher[1], Valentin Unger[1], Ulrike Hartmann[2], Jan Hochweber[1]
[1]Pädagogische Hochschule St.Gallen, Schweiz; [2]Leibniz-Institut für Bildungsforschung und Bildungsinformation Fraknfurt (DIPF)

Hintergrund & Fragestellung

Durch ihre Forschungsergebnisse soll Bildungsforschung eine wissenschaftliche Basis – Evidenz – zur Optimierung der Bildungsarbeit bereitstellen (z.B. Broekkamp & van Hout-Wolters, 2007). Forschungserkenntnisse finden jedoch nicht ohne Weiteres ihren Weg in die Schulen; Immer häufiger wird von einer "Kluft" zwischen Bildungsforschung und Praxis gesprochen (z.B. Farley-Ripple et al.,2018). Um diese Kluft überbrücken zu können, ist es wichtig, mehr über Lehrpersonen und ihr Verhältnis zur Nutzung von Forschungswissen herauszufinden. Denn anhand diesbezüglicher Erkenntnisse können Angebote für Lehrpersonen gezielt ausgebracht und gerahmt werden. Wie oder warum Forschungswissen in der Praxis (nicht) genutzt wird, ist bis heute noch nicht ausreichend beforscht (z.B. Cain, 2015). Fasst man die bestehende Literatur zur Nutzung von Forschungswissen zusammen (z.B. Demski, 2017, Hintzke, Gesang & Besa, 2020) scheinen sich über verschiedene Stichproben und Studienkontexte drei Gruppen von Lehrpersonen beschreiben zu lassen: Forschungskritische Lehrpersonen, die Forschungswissen überhaupt nicht nutzen (wollen), forschungsaffine Lehrpersonen, die bemüht sind Forschungswissen in ihren beruflichen Alltag einzubeziehen und eine «neutrale» Durchschnittsgruppe dazwischen. Einstellungen bzw. Überzeugungen hinsichtlich der Nutzung und der Relevanz von Forschungswissen wurden bei den einschlägigen Untersuchungen (z.B. Demski, 2017) nicht mit einbezogen. Es kann jedoch nicht davon ausgegangen werden, dass Einstellungen und tatsächliches Nutzungsverhalten zwangsweise korrespondieren (z.B. Hinke, 2020). Folgende Forschungsfrage wird überprüft: Lassen sich die drei Nutzungstypen von Lehrpersonen, wie aus der Literatur abgeleitet, in unserer Stichprobe replizieren und anhand von Einstellungen zum Forschungswissen genauer beschreiben?

Methode

Es steht eine Stichprobe von N = 123 Lehrpersonen zur Verfügung. Im Mittel haben die Lehrpersonen 13.48 (SD = 10.45) Jahre Berufserfahrung. 61.8% der Lehrpersonen stammen aus Deutschland, 35.0% aus der Schweiz. Die restlichen Lehrpersonen (3.2%) stammten aus Österreich. Je knapp die Hälfte der Lehrpersonen unterrichten an einer Primarstufe und an Sekundarschulen (jeweils 44.7%), der Rest der Lehrpersonen (10.6 %) an anderen Schulen. Zur Beantwortung der Forschungsfrage wurde eine latente Profilanalyse mit Mplus (Muthén & Muthén, 1998-2017) durchgeführt. Als Grundlage für die latente Profilanalyse dienten die Antworten der Lehrpersonen zur Nutzungshäufigkeit von Forschungswissen (1 Item), zur Nutzungsintention (5 Items) und Einschätzungen zur Relevanz von Forschungswissen (3 Items). Ergänzend wurden die vorhandenen Antworten der Lehrpersonen auf eine offene Zusatzfrage mittels qualitativer Inhaltsanalyse anhand der deduktiven Kategorien zu den theoretisch abgeleiteten Nutzungstypen ausgewertet. Für die Analyse der qualitativen Daten wurde MaxQDA (MAXQDA, 1989-2021) verwendet.

Ergebnisse

Für die Latente Profilanalyse wurden Modelle mit zwei und drei Klassen gerechnet und anhand verschiedener Modelkennwerte sowie inhaltlicher Sinnhaftigkeit der Profile verglichen. Weder die Modellkennwerte der 2-Profilelösung, noch die der 3-Profilelösung liefern Anlass, eines der Modelle abzulehnen. Zusammengefasst lässt sich jedoch eine leichte Überlegenheit für das Modell mit den drei Profilen festhalten. Die 3-Klassenlösung deckt sich weitestgehend mit den aus der Literatur abgeleiteten Nutzungstypen: Gruppe Forschungskritisch, Gruppe Forschungsdivers und Gruppe Forschungsaffin. Interessant ist die Beschreibung der grössten Gruppe um die forschungsdiversen Lehrpersonen (N=80). Hier wird die Relevanz von Forschungswissen für die Schulpraxis wie bei der Gruppe forschungsaffiner Lehrpersonen zwar anerkannt (M ~ 3.00), die Nutzungshäufigkeit von Forschungswissen (M=2.02) ist jedoch vergleichbar schwach ausgeprägt wie bei der Gruppe der Forschungskritischen Lehrpersonen. Die Gruppe der Forschungsdiversen Lehrpersonen äussert jedoch eine stärkere Intention Forschungswissen einzubinden (M=2.17). Auch aus der qualitativen Inhaltsanalyse ergibt sich Unterstützung für die Differenzierung der drei postulierten Nutzungstypen.

Diskussion

Insgesamt gesehen, geben Lehrpersonen an, selten Forschungswissen für die alltägliche Praxis zu nutzen. Die Ergebnisse der Profilanalyse zeigen aber auch, dass Nicht-Nutzung nicht automatisch bedeutet, dass Lehrpersonen Forschungswissen für unwichtig halten (Forschungsdiverse und Forschungsaffine Gruppe). Gleichzeitig liefern die Ergebnisse Hinweise darauf, dass Nicht-Nutzung von Forschungswissen nicht uneingeschränkt dadurch entgegengewirkt werden kann, dass seitens der Bildungsforschung Angebote geschaffen werden; Dann nämlich, wenn der Wunsch nach Einbezug von Forschungswissen nicht besteht (Forschungskritische Gruppe).

72. Entwicklung und Erprobung eines Instrumentes zur Erfassung der Strategien evidenzbasierte Informationen im schulischen Kontext online zu finden und zu bewerten

Jakob Meyer, Jennifer Paetsch
Otto-Friedrich-Universität Bamberg, Germany

Theoretischer Hintergrund

Digitale Kompetenzen von Lehrkräften erfahren seit einigen Jahren großes Forschungsinteresse. Dabei stehen etwa der Einsatz digitaler Lehrformen im Unterricht oder die Förderung des kompetenten Umgangs mit digitalen Medien im Fokus (Kultusministerkonferenz, 2016; Rubach & Lazarides, 2019). Eine weitere Komponente ist die Fähigkeit (angehender) Lehrkräfte, digitale und webbasierte Kanäle zu nutzen um Informationen zur Unterrichtsplanung und -gestaltung zu finden und zu bewerten (u.a. Trempler et al., 2015; Zimmermann et al., 2022).

Empirische Studie zeigen, dass Online-Ressourcen für Lehramtsstudierende mittlerweile die am häufigsten genutzte Informationsquelle sind (Fischer et al., 2022). Gleichzeitig sind digitale Kompetenzen unter Studierenden jedoch durchschnittlich nicht besonders stark ausgeprägt (Nagel et al., 2020). Es ist deshalb fraglich, ob (angehende) Lehrkräfte geeignete Strategien zum Suchen und Bewerten von Informationen im Internet einsetzen. Diese Fähigkeit ist relevant, da von Lehrkräften erwartet wird, ihren Unterricht gezielt und durch wissenschaftliche Erkenntnisse gestützt zu planen und durchzuführen (Kultusministerkonferenz, 2019). Um entsprechende empirische Inhalte zu finden und angemessen zu bewerten, scheinen spezifische Strategien und Kompetenzen notwendig zu sein.

Es liegen bereits einige Ansätze vor, um verschiedene Strategien der Suche und Bewertung voneinander differenzieren und zuverlässig erfassen zu können. Beispielsweise umfasst das Konstrukt des „Critical Online Reasoning“ (Molerov et al., 2020) Strategien der Suche von Informationen im Internet, die kritische Bewertung dieser Informationen sowie die Integration und der Abgleich mit bisherigem Wissen. Das „Critical Online Reasoning Assessment“, wie auch einige andere Instrumente zu Suchstrategien (u.a. Nagler et al., 2020) beziehen sich dabei auf allgemeine Rechercheaufträge. Für (angehende) Lehrkräfte ist die Suche nach evidenzbasierten Informationen jedoch besonders im pädagogischen Kontext bedeutsam. Hier setzt das vorliegende Vorhaben an, in dem ein Instrument entwickelt und erprobt wird, welches spezifisch die Strategien von Lehramtsstudierenden erfasst, evidenzbasierte Informationen online zu finden und zu bewerten.

Fragestellung

Ist das Instrument dazu geeignet, die Strategien von Lehramtsstudierenden bei der Online-Suche nach und Bewertung von evidenzbasierten Informationen für den Schulkontext valide und reliabel zu erfassen?

Methode

Für die Pilotierungsstudie werden Lehramtsstudierende der Otto-Friedrich-Universität Bamberg rekrutiert. Den Studierenden wird ein Fallbeispiel mit einer konkreten Rechercheaufgabe aus dem pädagogischem Kontext vorgelegt, welches in einem Computerlabor bearbeitet werden soll. Die Strategien werden einerseits mit einer adaptierten Version des „Online Information Searching Strategy Inventory“ (Tsai, 2009) erhoben. Zusätzlich werden Gedankenprotokolle und Nachgespräche geführt, um die verwendeten Strategien und die Kriterien, nach denen spezifische Quellen ausgewählt oder verworfen wurden, detaillierter zu erfassen. Neben den Strategien soll unter anderem auch auf den Aufbau des Instruments selbst, etwa die Eignung der Fallbeispiele, besprochen werden, um mögliche Probleme oder Verbesserungspotenziale offenzulegen.

Die Anzahl und Art besuchter Websites im Rahmen der Recherche wird mithilfe von Log-Daten protokolliert und ausgewertet. Außerdem werden Selbsteinschätzungen digitaler Kompetenzen (Rubach & Lazarides, 2019), Einstellungen zur Nützlichkeit empirischer Studien (Reinders et al., 2015) und Vorwissen zu empirischem Arbeiten als Kontrollvariablen bzw. zur Validierung erhoben.

Ergebnisse

Das Poster wird das entwickelte Instrument den Ablauf und die Ergebnisse der Pilotierungsstudie genauer vorstellen. Es wird dargestellt, ob und in welcher Form zuverlässig unterschiedliche Strategien zur Suche und Bewertung empirischer Quellen im Internet erhoben werden konnten. Außerdem werden Zusammenhänge zwischen den verwendeten Strategien sowie der Anzahl und Qualität verwendeter Quellen vorgestellt. Abschließend soll die Nützlichkeit des Instruments für Forschung und Lehrer*innenbildung diskutiert werden.

73. Online-Tools zur Förderung situationsspezifischer Kompetenzen (angehender) Lehrkräfte - Entwicklung und Evaluation

Katharina Frank, Olga Zlatkin-Troitschanskaia, Jasmin Reichert-Schlax, Christian Dormann, Sebastian Brückner
Johannes Gutenberg-Universität Mainz, Germany

In der Lehrkräfteausbildung sind Hochschulen dafür verantwortlich, professionelles Wissen von Lehramtsstudierenden zu fördern und ihnen situationsspezifische Kompetenzen zu vermitteln, damit sie in der komplexen Berufspraxis unterrichten können[1,2]. Damit Lehrkräfte in ihrem Studium hinreichend auf die beruflichen Herausforderungen vorbereitet werden[1,2], müssen sie ihr erworbenes Fachwissen in komplexen Unterrichtssituationen anwenden und in effektives Handeln transformieren können[3]. Daher ist in der Lehrkräfteausbildung der Einsatz hochwertiger Lehr-Lerntools unerlässlich, die die Komplexität der Unterrichtspraxis reflektieren und die situationsspezifischen Kompetenzen der Studierenden effektiv trainieren. In der zunehmend selbstgesteuerten Hochschullehre spielen digitale Lehr-Lerntools eine zentrale Rolle, da traditionelle Lehr-Lernansätze wirksam durch digitale Elemente ergänzt werden können[4]. Ziel des Einsatzes digitaler Tools ist die Gewährleistung besserer Lernergebnisse und höherer Gerechtigkeit[5], durch Flexibilität und Individualität in der Nutzung, um die Heterogenität der Studierenden zu adressieren[6]. Die Nachfrage nach frei zugänglichen Online-Tools, zur Förderung situationsspezifischer Lehrkompetenzen, wächst daher in der Lehrkräfteausbildung stetig[7,8,9].

Die für die Berufspraxis erforderlichen situationsspezifischen Lehrkompetenzen können in aktionsbezogene Kompetenzen (AK), die die spontane Bewältigung von Ereignissen im Unterricht fokussieren, und reflexive Kompetenzen (RK), um Unterricht in der Vor- und Nachbereitungsphase auszuwerten, unterteilt werden[10]. Basis für diese Lehrkompetenzen bietet das professionelle Wissen angehender Lehrkräfte, das sich in die Facetten content knowledge (CK), pedagogical content knowledge (PCK) und pedagogical knowledge (PK)[11] sowie technological pedagogical content knowledge (TPACK) unterteilt[12]. AK und RK werden bei der praktischen Bewältigung komplexer Situationen unter Verwendung von CK, PCK, PK und TPACK angewandt[10,11].

Trotz hoher Nachfrage sind bisher nur wenige qualitativ hochwertige Online-Tools in der Lehrkräfteausbildung frei verfügbar, die eine Förderung situationsspezifischer Lehrkompetenzen adressieren. In einem vorangegangenen Projekt wurde daher ein digitales Tool zur Förderung von AK und RK angehender Wirtschaftslehrkräfte entwickelt, validiert[13,14,15], dessen Wirksamkeit in einer Prä-Post-Interventionsstudie im Kontrollgruppendesign belegt[16] und als Open Educational Ressource (OER) frei zugänglich hochgeladen. Im Folgeprojekt werden Tools entwickelt und evaluiert, die die professionelle Wissensbasis mit den Facetten PCK (im Fach Wirtschaft), PK (u.a. selbstreguliertes Lernen, Gesprächsführung im Unterricht[17,18,19]) sowie TPACK (u.a. Modelle digitaler Lehre ICAP/SAMR, Aufbau und Funktion von Lernmanagementsystemen[20,21,22]) fokussieren und ebenfalls auf einer OER-Plattform hochgeladen werden. Um die Multidimensionalität einer Unterrichtssituation abzubilden, berücksichtigen die Tools Prinzipien der Problemorientierung und Selbststeuerung und beinhalten Schriftdateien zur Beschreibung der Lehrkontexte und -ziele sowie Wissens- und Kompetenzbeschreibungen, instruktionale Videovignetten zur authentischen Darstellung der Berufspraxis[23,24,25,26,27], audiovisuelle Medien wie Erklärvideos, computergestützte Lernaufgaben mit Instruktions- und Feedback-bezogenen Elementen[1] sowie Reflexions- und (Selbst-)Kontrollelemente[28].

In diesem Projekt wurde bereits die Wirksamkeit der Tools mit Schwerpunkt auf PK und TPACK (N=250) in einer Prä-Post-Interventionsstudie im Kontrollgruppendesign mit folgender

Forschungsfrage untersucht: *Führt das Lernen mit den Online-Tools zu einer Steigerung der geförderten Wissensfacetten im Posttest (t2) im Vergleich zum Pretest (t1) in der Interventions- im Vergleich zur Kontrollgruppe?*

Erste Ergebnisse zeigen signifikante Entwicklungen in den Wissensfacetten PK Gesprächsführung in der IG (N=46) (x_{t1}=2.98±1.341, x_{t2}=3.65±1.286, t(45)=-3.065, p=0.004, d=0.44) und keine bedeutsamen Entwicklungen in der KG (N=19) (x_{t1}=3.47±1.541, x_{t2}=3.05±1.393, t(18)=-1.140, p=0.269, d=0.25) sowie PK selbstreguliertes Lernen in der IG (N=34) (x_{t1}=5.65±0.692, x_{t2}=6.02±0.620, t(33)=-4.573, p=0.000, d=0.75) und keine bedeutsame Entwicklung in der KG (N=82) (x_{t1}=1.72±0.972, x_{t2}=1.80±1.082, t(81)=-0.757, p=0.451, d=0.089) sowie TPACK ICAP/SAMR/LMS in der IG (N=55) (x_{t1}=8.11±2.580, x_{t2}=10.09±3.129, t(54)=-5.009, p=0.000, d=0.71) und keine bedeutsame Entwicklung in der KG (N=14) (x_{t1}=8.43±2.344, x_{t2}=8.64±2.951, t(13)=-0.233, p=0.590, d=0.08).

Die Ergebnisse deuten darauf hin, dass die Online-Tools wirksam zu einer Verbesserung der Lehrkräfteausbildung beitragen können. Weitere Tools werden im kommenden Wintersemester 2022/23 an vier Hochschulen bundesweit und international, zwei Studienseminaren und einer berufsbegleitenden Lehrerfortbildung eingesetzt und evaluiert. Die Überarbeitung der evaluierten Tools und die Evaluierung weiterer Tools (mit Fokus auf PCK) sind die nächsten Schritte. Alle Tools werden frei zugänglich als OER hochgeladen. Die Ergebnisse dieser Studie werden im Hinblick auf Limitationen und Implikationen für eine systematischere Ausrichtung der Lehrkräfteausbildung diskutiert.

74. Situationsbezogene mathematikdidaktische Kompetenzen und fachlich-fachdidaktische Überzeugungen von Primarstudierenden und ihren Praxislehrpersonen im Vergleich

Esther Brunner[1], Sanja Stankovic[1], Marco Galle[2], Sonja Hiebler[2], Annelies Kreis[2]
[1]Pädagogische Hochschule Thurgau, Kreuzlingen (Schweiz), Switzerland; [2]Pädagogische Hochschule Luzern, Luzern (Schweiz)

1. Theoretischer Hintergrund

Die Ausbildung von Primarlehrpersonen in der Schweiz erfolgt an Pädagogischen Hochschulen als dreijähriges Bachelor-Studium, das zur Lehrbefähigung in mindestens sieben Unterrichtsfächern führt. Für die einzelnen Fächer steht daher nur eine geringe Stundendotation zur Verfügung, um eine fachlich-fachdidaktische Grundausbildung zu ermöglichen. Für die Mathematikausbildung sind dafür an den verschiedenen Schweizer Hochschulen total zwischen 7 und 12 ECTS für den obligatorischen Bereich vorgesehen (z. B. PH Luzern, 2021; PHTG, 2017). In dieser Zeit sollen nebst curricularem Grundwissen (MCK) auch fachliche diagnostische Kompetenzen (MPCK: DK) sowie Kompetenzen zur fachlichen Unterrichtsgestaltung (MPCK: UGK) (Baumert & Kunter, 2011, 2013) vermittelt werden.

Ergänzt wird das Lehrangebot an den Hochschulen durch verschiedene mehrwöchige Praktika in Primarklassen, in denen die Studierenden unter Anleitung der Klassenlehrpersonen (Praxislehrpersonen) unterrichten. Diese berufspraktische Ausbildung umfasst ca. 25 % der Studienzeit (EDK, 2015). Dadurch sollen u.a. fachlich-fachdidaktische Inhalte, die an der Hochschule vermittelt werden, in der Praxis erprobt und anschliessend reflektiert werden.

Damit diese enge Verzahnung von Lernen an der Hochschule und Lernen im Praktikum produktiv erfolgen kann, ist es notwendig, dass die Studierenden in den Feldern Hochschule und Praxisklasse ein kohärentes Verständnis von (fachlichem) Lernen vorfinden und dass sie in der Praxis auf Lehrpersonen mit fachlich-fachdidaktischer Expertise stossen. Es ist daher zu erwarten, dass die Praxislehrpersonen als erfahrene Lehrpersonen ein Mehr an situationsbezogenem fachlich-fachdidaktischem Wissen aufweisen, in den fachlichen Überzeugungen aber den Studierenden ähneln. Im Rahmen des Beitrags interessiert daher folgende Fragestellung: Inwiefern zeigen sich bezüglich situationsbezogenen mathematikdidaktischen Kompetenzen und fachlichen Überzeugungen Gemeinsamkeiten und Unterschiede zwischen Primarstudierenden und Praxislehrpersonen?

2. Methode

Im interdisziplinären SNF-Projekt «Lerngelegenheiten für Lehrstudierende im sozialen Netzwerk Praxisfeld aus allgemein- und mathematikdidaktischer Perspektive» (Kreis & Brunner, 2019, im Druck) wurden bei einer ersten (von insgesamt zwei geplanten) Kohorte von Primarstudierenden (N = 186) und Praxislehrpersonen (N = 257) mathematikdidaktische Kompetenzen situationsbezogen anhand von Vignetten erfasst (Stankovic & Brunner, 2022). Die acht Vignetten erfassten jeweils anhand einer Unterrichtssituation mit je vier Items mathematikspezifisches diagnostisches Wissen (DK) und mathematikspezifisches Unterrichtsgestaltungs- bzw. Unterstützungswissen (UGK). Zudem wurden Beliefs zu Mathematik und zu Lehren und Lernen von Mathematik erhoben (Laschke & Felbrich, 2014). Die Skalen zum Fach beschreiben eine Regel- und Prozedurorientierung und eine Sicht von Mathematik als Prozess des Entdeckens und Erforschens. Lehren und Lernen umfasst eine konstruktivistisch und eine transmissiv geprägte Sicht.

Die Daten wurden metrisch erfasst und als Mittelwerte (0-1) für die diagnostische und die Unterrichtsgestaltungskompetenz sowie als Mittelwert pro Beliefskala (1-6) berechnet und bezüglich systematischer Gruppenunterschieden zwischen den Expert*innen und Noviz*innen mittels t-Tests analysiert.

3. Ergebnisse und ihre Bedeutung

Die situationsbezogenen mathematikdidaktischen Kompetenzen von Praxislehrpersonen und Primarstudierenden unterscheiden sich nicht bezüglich diagnostischer Kompetenz, sondern nur bezüglich Unterrichtsgestaltungskompetenzen (t = 4.15; df = 370.88, inhom. Varianzen korr., p < .001). Diese fallen bei den Praxislehrpersonen leicht höher aus als bei den Studierenden (M_{PraL} = .71; SD = .12; M_{Stud} = .65; SD = .16).

Bei der Sicht von Mathematik als Prozess von Erforschen zeigen die Praxislehrpersonen ebenfalls höhere Werte (M_{PraL} = 4.93, SD = .57; M_{Stud} = 4.45, SD = .80) als die Studierenden (t = 7.39; df = 440.99, inhom. Varianzen korr., p < .001) bei einer vergleichbar moderaten Sicht von Mathematik als Regel und Prozedur. Zum Lernen und Lehren zeigen beide Gruppen vergleichbare niedrige Werte in der transmissiven Skala und sehr hohe bei der konstruktivistischen, die bei den Praxislehrepersonen signifikant stärker ausgeprägt sind als bei den Studierenden (t = 2.55; df = 437, p < .01).

Die Befunde werfen aufgrund der Ähnlichkeit fachlich-fachdidaktischer Überzeugungen und diagnostischer Kompetenzen von Studierenden und Praxislehrpersonen die Frage auf, inwiefern aus fachlich-fachdidaktischer Sicht von Expert*innen und Noviz*innen ausgegangen und der Ausbildungsauftrag der Praxis auch in fachlicher Hinsicht erfüllt werden kann.

75. Leistungsbasierte Bewertung der diagnostischen Kompetenz von Lehramtsstudent:innen in Bezug auf die Lernaktivitäten von Schüler:innen: Validierung einer Simulation

Meral Roeben
LMU München - Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie (Prof. Fischer)

Eine der zentralen Kompetenzen von Lehrkräften ist es, Schüler:innen optimal zu aktivieren, um Lernaktivitäten und Lernprozesse herbeizuführen, die sich an den angestrebten Lehr- und Bildungszielen orientieren. Heute schließt das unweigerlich auch die Aktivierung durch digitale Medien mit ein. Es sollen also effiziente und erfolgreiche Lernprozesse angeregt und ermöglicht werden - erstens durch eine Unterrichtsplanung, in der Lernziele und Lernaktivierung systematisch aufeinander bezogen werden und zweitens durch eine Lernbegleitung der Schüler:innen im Unterricht selbst, welche diese bedarfsgerecht bei ihren Lernaktivitäten unterstützt. Beides setzt insbesondere auch diagnostische Kompetenzen in der Planungs- und Durchführungsphase des Unterrichts voraus. Chi zeigt, dass eine optimale Passung von Lernziel und Aktivierung zu besseren Lernprozessen führen kann. Bei manchen Lernzielen kann eine passive Aktivierung ökonomisch sein, anspruchsvollere Lernziele verlangen nach konstruktiver oder interaktiver Aktivierung. Ob Lehrer:innen über die entsprechenden Kompetenzen verfügen, wurde bisher vor allem anhand von Selbsteinschätzungen gemessen und ist mit den bekannten Problemen hinsichtlich der Zuverlässigkeit der damit gemessenen Kompetenzen behaftet. Um dem entgegenzuwirken, wurde im Rahmen eines Projektes des Kompetenznetzwerks Medienbildung und Digitalisierung (KMBD) der LMU eine Simulation entwickelt. Als „approximation of practice“ ist die Simulation ein objektives Testinstrument und misst die Diagnosekompetenzen von Lehrer:innen bei der Unterrichtsplanung und -durchführung. Die vorliegende Studie validiert diese Simulation.

Die Simulation ist als „Point-and-Klick"-Adventure umgesetzt. Die teilnehmenden Lehramtsstudent:innen schlüpfen dabei in die Rolle von Referendar:innen an einer weiterführenden Schule, wo sie von ihrer Seminarlehrerin empfangen und instruiert werden. In einer ersten Phase analysieren und beurteilen sie Unterrichtsentwürfe hinsichtlich deren Aktivierungspotential. Im Mittelpunkt der Diagnose steht die Frage, wie gut die geplante Aktivierung der Schüler:innen den Lernzielen der jeweiligen Unterrichtssequenz entspricht. Die Lernaktivitäten der Schüler:innen sind dabei gemäß dem ICAP-Modell von Chi operationalisiert. Nachdem die Proband:innen (in ihrer Rolle als Referendar:innen) gemeinsam mit ihrer Seminarlehrerin optimale Unterrichtsentwürfe gemäß dem ICAP-Modell identifiziert und ausgewählt haben, begeben sie sich zur Hospitation in den Unterricht unterschiedlicher Klassen. Dort erfolgt eine Aktivierung der Schüler:innen auf Grundlage der ausgewählten Unterrichtsentwürfe. Die Simulationsteilnehmer:innen beobachten die ausgeführten Lernaktivitäten der einzelnen Schüler:innen. Ihre Aufgabe ist es nun, zu bewerten, auf welcher ICAP-Stufe die Schüler:innen aktiviert sind und ob eine Unterstützung durch die Lehrkraft erforderlich ist, um die Schüler:innen optimaler zu aktivieren.

Um die Leistungen der Proband:innen in der Test- und Lernumgebung der Simulation auswerten und interpretieren zu können, ist der Simulation ein Test zum allgemeinen Aktivierungswissen vorgeschaltet. Außerdem müssen die Teilnehmer:innen nach dem ersten Fall der Planungsphase Wissensfragen zum ICAP-Modell beantworten. Zusätzlich werden der intrinsische Cognitive Load sowie die Präsenz und das Authentizitätsempfinden erhoben. Dies ermöglicht eine genaue Einschätzung der Leistungen der Student:innen in der Simulation hinsichtlich ihrer Validität sowie des Zusammenhangs zwischen diagnostischen Kompetenzen sowie allgemein-didaktischem und spezifischem Vorwissen.

Es ergeben sich folgende Forschungsfragen:

1. Verfügen die Proband:innen, die besser beim allgemeinen Aktivierungstest abschneiden, über höhere Diagnosekompetenzen (repräsentiert durch die Punktescores aus den Fällen der Planungs- und Realisierungsphase)?

2. Verfügen die Proband:innen, die besser beim ICAP-Wissenstest abschneiden, über höhere Diagnosekompetenzen?

3. Ist der intrinsische Cognitive Load der Proband:innen niedriger, wenn sie über mehr Vorwissen (allgemeines Aktivierungswissen und ICAP-Wissen) verfügen?

4. Ist der intrinsische Cognitive Load der Proband:innen niedriger, wenn die Diagnosekompetenzen höher sind?

Die Simulation gilt als validiert, wenn die Proband:innen mit höherem Vorwissen (allgemeines Aktivierungswissen und ICAP-Wissen) über ausgeprägtere Diagnosekompetenzen verfügen und gleichzeitig einen niedrigeren intrinsischen Cognitive Load aufweisen.

Nachdem die Simulation entwickelt und auf der Basis eines Cognitive Labs mit sieben Teilnehmer:innen angepasst wurde, wird derzeit eine Validierungsstudie zur Simulation durchgeführt. Dabei wird eine Stichprobe von N=140 angestrebt. Die Datenerhebung soll bis November 2022 abgeschlossen und auf der Konferenz präsentiert werden. In ersten Auswertungen auf der Grundlage der aktuellen Stichprobe von N=74 deuten sich positive Befunde an, welche eine erfolgreiche Validierung der Simulation versprechen.

76. Lernverlaufsdiagnostik zur Förderung schulischer Fähigkeiten? Ein metaanalytischer Review zur Effektivität.

Amelie Fuchs, Anika Radkowitsch, Daniel Sommerhoff
IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany

Individualisierter, an Schüler:innen orientierter Unterricht, bedarf einer andauernden Erfassung des Lernstandes und individuellen Lernfortschritts (Jung et al., 2018). Diese sollte reliabel und valide, einfach zu handhaben, ohne großen Aufwand implementierbar, auf die Lehrinhalte abgestimmt (Lembke & Stecker, 2007) und zwischen Messzeitpunkten vergleichbar sein. Ein erfolgsversprechender Ansatz, bei dem individuelle Lernverläufe über die Zeit abgebildet werden, ist die Lernverlaufsdiagnostik. Lernverlaufsdiagnostik beinhaltet wiederholte, kurze Tests desselben Inhalts, mit denen die individuellen Fortschritte von Schüler:innen gemessen und an Lehrkräfte (und ggf. Schüler:innen) zurückgemeldet werden können (Deno, 2003; Fuchs, 2004). Im Sinne einer

formativen Beurteilung (Black & William, 2009) soll so eine effiziente, individuelle Förderung der Schüler:innen ermöglicht werden, da positive wie negative Leistungsentwicklungen zeitnah sichtbar werden und Lehrkräfte wie ggf. Schüler:innen sich daran orientieren können. Trotz des Ursprungs der Lernverlaufsdiagnostik aus der Sonderpädagogik mit Fokus auf sprachlichen Fähigkeiten, eignet sie sich prinzipiell auch zur Erfassung weiterer Fähigkeiten (z.B. mathematische Fähigkeiten, Foegen et al., 2007) im Regelschulbereich und ist durch die fortschreitende Digitalisierung zunehmend einfacher implementierbar (Souvignier, 2018).

Studien zur Güte von Instrumenten der Lernverlaufsdiagnostik zeigen, dass es Instrumente gibt, mit denen eine valide und reliable Abbildung des Lernstandes (z.B. Romig et al., 2017; Yeo, 2011) sowie eine sensitive Erfassung der Leistung von Schüler:innen über die Zeit (z.B. Wayman et al., 2007) möglich ist. Bisher gibt es jedoch wenige Studien zur Effektivität von Lernverlaufsdiagnostik, welche aufgrund unterschiedlicher Settings und Umsetzungen ein heterogenes und eher unsystematisches Forschungsbild erzeugen. Zwar deuten erst Ergebnisse darauf hin, dass Lernverlaufsdiagnostik sich positiv auf die Leistung von Schüler:innen auswirken kann, gleichermaßen zeigt sich hier eine hohe Heterogenität (Fuchs & Fuchs, 1986; Stecker et al., 2005). So scheint Lernverlaufsdiagnostik unter verschiedenen Bedingungen (z.B. Adaption des Unterrichts, Stecker et al., 2005: Visuelle Aufbereitung der Leistungsrückmeldung, Fuchs & Fuchs, 1986) unterschiedlich effektiv. Vor diesem Hintergrund vermag eine Systematisierung bisheriger Ergebnisse durch ein Literaturreview sinnvoll und richtungsweisend für zukünftige Forschung und Praxis im Bereich der Lernverlaufsdiagnostik zu sein.

Der präsentierte systematische Review trägt daher den Forschungsstand zur Effektivität von Lernverlaufsdiagnostik zusammen und betrachtet diesen metaanalytisch. Hierbei werden die folgenden Fragestellungen untersucht: (i) Welchen Einfluss hat Lernverlaufsdiagnostik auf den Erwerb schulischer Fähigkeiten von Schüler:innen? (ii) Inwiefern variiert die Effektivität der Lernverlaufsdiagnostik basierend auf Eigenschaften der Stichprobe, Eigenschaften der Lernverlaufsdiagnostik, begleitenden Maßnahmen bzw. dem Fähigkeitsbereich?

Zur Beantwortung der Fragestellungen wurden zwei systematische Literaturdatenbankensuchen durchgeführt, wobei initial der Fähigkeitsbereich Mathematik fokussiert und in der zweiten Suche auf beliebige Bereiche erweitert wurde. Eingeschlossen wurden Studien, die die Effektivität von Lernverlaufsdiagnostik in einem (quasi-)experimentellen Design untersuchten und in einer begutachteten englisch-sprachigen Zeitschrift veröffentlicht wurden. Für den Bereich Mathematik wurden 11 geeignete Studien (κ = .70) mit insgesamt 3277 untersuchten Schüler:innen, 37 Gruppenvergleichen sowie 17 unabhängigen Effekten identifiziert. Diese Studien wurden anhand eines deduktiv-induktiv entwickelten Kodierschema kodiert (Cohens Kappa κ_{mean} = .92 [.70; 1]). Die Auswertung der erweiterten Suche über den Fähigkeitsbereich Mathematik hinaus laufen aktuell.

Der Haupteffekt von Lernverlaufsdiagnostik auf die betrachteten mathematischen Fähigkeiten von Schüler:innen ist g = .54 (p < .001). Mittels 3PSM und trim-and-fill konnte kein Hinweis auf einen Publication bias festgestellt werden. Die vorläufigen Ergebnisse für den Bereich Mathematik zeigen, dass bisherige Studien vor allem aus dem Bereich Grundschule und Sekundarstufe I stammen.

Zudem wurden in den meisten Untersuchungen die Lehrkräfte aufgefordert den Unterricht, basierend auf den Ergebnissen der Lernverlaufsdiagnostik zu adaptieren. Neben der Lehrkraft erhielten überwiegend auch die Schüler:innen eine Leistungsrückmeldung.

In den bisher betrachteten Studien wird ein mittlerer positiver Effekt von Lernverlaufsdiagnostik auf die mathematischen Fähigkeiten von Schüler:innen deutlich. Die vorläufigen Ergebnisse des Reviews zeigen weiter, dass bezüglich der Untersuchung spezifischer Wirkmechanismen von Lernverlaufsdiagnostik sowie Einflussfaktoren auf diese weiterer Forschungsbedarf besteht.

77. Teacher emotions and their linkages with student outcomes in secondary education. A multilevel modeling approach

Marina Elena Pfeifer[1], Shruti Athavale[2], Anne Christiane Frenzel[1]
[1]Ludwig-Maximilians-Universität München, Deutschland; [2]Technische Universität München, Deutschland

The role of teacher well-being is critical to students' school life. A model recently proposed by Frenzel and colleagues (2021) describes how teacher emotions are linked with student outcomes via teacher behaviors such as relationship building, classroom management, and instructional strategies. There is extensive literature linking teacher emotions to student competence beliefs (Mueller & Dweck, 1998), motivation (Becker et al., 2014; Frenzel et al., 2009), and performance (Banerjee et al., 2017; Frenzel et al., 2020). Consistent with the model, research suggests that teacher joy and anger impact three core criteria for qualitative teaching, namely a supportive climate, classroom management, and cognitive activation (Becker et al., 2015; Chen, 2018; Hagenauer et al., 2015; Keller et al., 2014; Klieme et al., 2009). These criteria have been associated with students' greater competence beliefs, motivation, and performance (Brophy, 2006; Hamre & Pianta, 2005).

The current study sought to validate parts of the model by Frenzel et al. (2021). We used existing data from 17,554 students nested within 679 teacher clusters from the Global Teaching InSights longitudinal study (OECD, 2018). This study included teacher-rated joy and anger and student-rated teacher behaviors, as well as data on student performance, self-reported competence beliefs, and motivation. Using multilevel structural equation modeling, we analyzed the effects of teacher joy and anger on student competence beliefs, motivation, and performance, with student ratings of teacher-student relationships, classroom management, and cognitive activation strategies serving as mediators.

Our results showed moderate to good fit for the joy and anger models, with both models fitting equally well, i.e., CFI = .93, TLI = .92, RMSEA = .04, SRMR = .03, and χ^2-p < .001. Teacher joy indirectly positively affected student competence beliefs via classroom management (.16**) and cognitive activation (.15***), as well as student motivation via teacher-student relationship (.12**) and classroom management (.20**). In addition, teacher joy had a positive indirect effect on student performance via classroom management (.30*) and, contrary to expectations, a negative indirect effect via cognitive activation (−.23***), which was qualified by a positive link between teacher joy and cognitive activation, and negative link between cognitive activation and performance. Further, teacher anger had a negative indirect effect on student competence beliefs via the teacher-student relationship (−.09**), classroom management (−.20***), and cognitive activation (−.19**), as well as on student motivation via the teacher-student relationship (−.19***) and classroom management (−.24***). Lastly,

teacher anger had a negative indirect effect on student performance via relationship (−.36***) and classroom management (−.42***).

We conclude from our results that teacher emotions are critically important to students insofar as they are systematically associated with the quality of teacher behaviors and in consequence, with student outcomes. As proposed in the model, and also recognizing that the data presented here are essentially cross-sectional, there is the possibility that reciprocal effects emanate from students back on teacher emotions via their appraisals (Frenzel et al., 2021). Thus, our findings are broadly aligned with previous research and reemphasize the importance of teacher well-being. Further investigations on the order of effects are required. Limitations and implications regarding teacher emotions and effective teaching are discussed.

78. Emotionen von Lehrpersonen und deren Auslösebedingungen im Teamteaching-Unterricht

Franziska Mühlbacher[1], Gerda Hagenauer[1], Melanie M. Keller[2]
[1]University of Salzburg, Austria; [2]IPN Kiel - Leibniz Institute for Science and Mathematics Education

Lehrpersonen erleben im Unterricht eine Vielzahl an Emotionen, wie beispielsweise Freude oder Ärger (Frenzel, 2014; Hascher & Krapp, 2014), die in Zusammenhang mit der Instruktionsqualität, Lehrer*innen-Schüler*innen Beziehung oder dem Wohlbefinden der Lehrperson stehen (Chen, 2018; Hagenauer, Hascher & Volet, 2015; Keller et al., 2014). Als emotionsauslösende Quelle von Lehrer*innenemotionen standen bislang hauptsächlich Schüler*innen im Forschungsfokus (Frenzel et al., 2009). Im Gegensatz zum „traditionellen“ Unterricht, bei dem eine Lehrperson unterrichtet, unterrichten beim kooperativen Unterrichtssetting des Teamteachings Lehrpersonen gemeinsam und teilen sich die Verantwortung für die Planung, Durchführung und Evaluation des Fachunterrichts in einer Klasse (Krammer et al., 2017). Basierend auf appraisal-theoretischen Annahmen zur Emotionsentstehung (Frenzel, 2014) ist es naheliegend, dass Teamteaching-Lehrpersonen im Teamunterricht Emotionen nicht nur aufgrund der Schüler*innen, sondern auch aufgrund der zweiten Lehrperson erleben, da auch diese Interaktionssituationen im Hinblick auf die zugrunde liegenden Ziele und deren Erreichung oder Nicht-Erreichung bewertet werden.

Daher wird in dieser Studie die Zielsetzung verfolgt, das Emotionserleben von Teamteaching-Lehrpersonen, das in der Interaktion mit der Partnerlehrperson entsteht, sowie die Auslösefaktoren dieser Emotionen zu untersuchen. Es nahmen 30 österreichische Teamteaching-Lehrpersonen an einer qualitativen, leitfadenbasierten Interviewstudie teil (23 weiblich, 7 männlich). Die Interviews (mittlere Dauer: 70 Minuten) wurden wortwörtlich transkribiert und anhand der strukturierenden, qualitativen Inhaltsanalyse (Mayring, 2015) ausgewertet. Die Intercoderreliabilität ist zufriedenstellend (korr. Kappa > .75).

Die Ergebnisse zeigen, dass das Emotionserleben von Teamteaching-Lehrpersonen und die Auslösefaktoren im Teamunterricht vielfältig sind. Insgesamt berichteten Lehrpersonen 47 diskrete Emotionen (22 positive, z. B. Freude, Heiterkeit, Glück; und 25 negative, z.B. Ärger, Langeweile, Unsicherheit).

Bezüglich der Auslösebedingungen konnten acht Hauptkategorien gebildet werden, welche wiederum 37 Subkategorien beinhalten. Konkret berichteten die Lehrpersonen, dass eine Übereinkunft im Team

bezüglich *gemeinsamer Ziele und Werte* relevant für deren positive Emotionen war; eine fehlende Übereinkunft dahingegen negative Emotionen auslösen konnte. Des Weiteren nannten Teamteaching-Lehrpersonen, dass das *Teamteaching-Setting an sich*, d.h. die Anwesenheit der zweiten Lehrperson in der Klasse, Emotionen wie Entspannung, aber auch Angst und Stress auslösen kann. *Charakteristika der Zusammenarbeit*, wie geteilte Planung und Vorbereitung, ein grundlegendes Interesse an der Zusammenarbeit, Verlässlichkeit und das Einhalten von Abmachungen, gegenseitige (fachliche und nicht-fachliche) Unterstützung oder eine faire Arbeitsteilung lösten ebenso positive Emotionen aus, während das Nicht-Erfüllen dieser Charakteristika negative Emotionen hervorrief. Laut Teamteaching-Lehrpersonen sind auch *Charakteristika bezogen auf das Unterrichtsverhalten des/der Teampartnerin*, wie die Offenheit und Kompromissbereitschaft der Partnerlehrperson, deren Methodenvielfalt und Kreativität, Flexibilität und Spontanität oder der Unterrichtsstil emotional bedeutsam. Viele Lehrpersonen erachteten es zudem als wichtig, dass sie ihr Potenzial in den Teamunterricht durch die Einnahme einer aktiven Rolle einbringen können. Des Weiteren berichteten die Lehrpersonen, dass eine ungewollt passive Rolle im Teamunterricht zu Ärger, Unmut oder Traurigkeit führen kann. *Charakteristika der Partnerlehrperson*, wie deren Einstellung zu und Einhaltung von schulinternen Teamteaching-Regelungen, sind ebenfalls emotional relevant. Lehrpersonen berichteten, dass auch *persönliche Merkmale*, wie der eigene Leistungsanspruch im Rahmen des Teamteachings, Angst auslösen. Eine wichtige Quelle von Emotionen liegt ebenso im *Umgang der Partnerlehrperson mit Schüler*innen*, beispielsweise in der Form von (un)fairer Behandlung von Schüler*innen oder Lob. Letztlich lösen auch *Aspekte innerhalb der interpersonalen und professionellen Beziehung der beiden Partnerlehrpersonen*, wie Wertschätzung und Respekt, Ehrlichkeit oder Loyalität, aber auch Unterbrechungen, unterschiedlichste Emotionen aus.

Die Ergebnisse der Studie zeigen, dass die Praxis des Teamteachings für Teamteaching-Lehrpersonen emotional bedeutsam ist und dass insbesondere das professionelle und interpersonale Verhalten des Teampartners/der Teampartnerin – bezogen auf die Zusammenarbeit aber auch auf die Umsetzung von Unterricht und den Umgang mit den Schüler*innen – eine wesentliche Quelle des Emotionserlebens sind. Für die Professionalisierung von Lehrpersonen im Teamteachingunterricht unter einer emotionalen Perspektive sind folglich klare Erwartungen an das Teamteaching sowie soziale Kompetenzen aber auch hohe Professionalität im Unterrichtshandeln von Bedeutung.

79. Der Quereinstiegsmaster als nicht-grundständiger Professionalisierungsweg für das Lehramt – Einordnung, Befunde, Perspektiven

Novid Ghassemi Tabrizi, Volkhard Nordmeier
Freie Universität Berlin, Deutschland

Nicht-grundständige Wege in den Lehrer*innenberuf wurden in den vergangenen Jahren im Kontext des Mangels an Lehramtsabsolvent*innen für unterschiedliche Fächer und Schulformen in nahezu allen Bundesländern geschaffen (Abs, Kuper & Martini, 2020; Brinkmann & Müller, 2020). Aktuelle Modellrechnungen deuten weiterhin auf die Notwendigkeit des Anbietens alternativer Wege in den Beruf in den kommenden Jahren hin (Klemm, 2021, 2022). Um auch die Qualität alternativer Wege in den Lehrer*innenberuf gewährleisten zu können, bedarf es sowohl der Einführung von (Mindest-

)Standards (GFD, 2018; Terhart, 2020)) als auch empirischer Forschung zu den professionellen Handlungskompetenzen der über alternative Wege qualifizierten Lehrer*innen (Richter, Pech & Zorn, 2021).

Der alternative Weg des *Quereinstiegs während des Studiums* unterscheidet sich hinsichtlich des zeitlichen Umfangs, der Lehrinhalte und der Einbindung lehrer*innenbildender Universitäten grundsätzlich gegenüber anderen nicht-grundständigen Wegen in den Lehrer*innenberuf. Quereinstiegsstudiengänge für das Lehramt (kurz ‚Q-Master') wurden unter anderem an Berliner Universitäten in den vergangenen Jahren eingerichtet. An der Freien Universität Berlin (Q-Master Integrierte Sekundarschulen/Gymnasien) und an der Humboldt-Universität zu Berlin (Q-Master Grundschulpädagogik) werden sie durch Forschungsprojekte begleitet (Ghassemi, Milster & Nordmeier, 2019; Lucksnat, Fehrmann, Müncher, Pech & Richter, 2022).

An der FU Berlin wurden im Zuge der Begleitforschung Studierende des Q-Masterstudiengangs und des regulären Lehramtsmasters seit dem Jahr 2016 mittels Tests und Fragebögen befragt: So wurden die Q-Masterstudierenden (n = 167) unter Anderem zu ihren *Berufswahlmotiven* (Watt & Richardson, 2007; König & Rothland, 2012), *Überzeugungen zum Lehren und Lernen* (Strauß, König & Nold, 2019), *Selbstwirksamkeit* (Pfitzner-Eden, Thiel & Horsley, 2014) und pädagogischen Vorerfahrungen befragt.

Im Studienfach Physik wurde außerdem exemplarisch die Entwicklung fachlicher Aspekte professioneller Handlungskompetenz (*fachdidaktisches Wissen, Fachwissen* und *Überzeugungen zum Lehren und Lernen im Fach*) zu Beginn und Ende des Masterstudiums mittels Tests und Fragebögen (Riese, 2009; Lamprecht, 2011; Korneck, Krüger & Szogs, 2017) längsschnittlich erfasst und mit Studierenden des regulären Lehramtsmasters an der FU Berlin verglichen. Ergänzt wird dieses Vorgehen durch leitfadengestützte Interviews zu zwei Befragungszeitpunkten. Ziel der Interviews ist es, weitere Daten zur formativen Evaluation des Modellstudiengangs zu sammeln und individuelle Entscheidungsprozesse, persönliche Ziele und den erlebten Nutzen für die Professionalisierungstrajektorien der über diesen alternativen Weg für den Lehrer*innenberuf qualifizierten Personengruppe nachzuzeichnen und zu vergleichen.

Die Auswertung der quantitative Begleitforschung an der FU Berlin deutet auf eine Ausprägung und Entwicklung professioneller Kompetenzen hin, welche mit Studierenden des regulären Lehramtsmaster vergleichbar ist (Ghassemi & Nordmeier, 2021a, 2021b). Bezüglich der Motive zur Berufswahl scheinen die Q-Masterstudierenden vornehmlich intrinsisch-altruistisch motiviert und damit mit jenen des regulären Lehramts vergleichbar (Ghassemi & Nordmeier, 2021b). Aufgrund der noch niedrigen Zahl Q-Masterstudierender pro Jahrgang im Fach Physik sind die vorgenommenen Hypothesentests zwar hinsichtlich der erreichbaren Teststärke limitiert, zugleich werden die Ergebnisse aber durch die Begleitforschung zum Q-Maser für das Grundschullehramt der HU Berlin größtenteils bestätigt (Lucksnat et al., 2022).

Die Auswertung der Interviewdaten mittels qualitativer Inhaltsanalyse nach Kuckartz (2018) deutet außerdem darauf hin, dass weniger geradlinige Berufsbiografien nicht nur bei Q-Masterstudierenden vorkommen. Für die persönliche Professionalisierung werden insbesondere das Praxissemester

sowie fachdidaktische und bildungswissenschaftliche Lehrveranstaltungen als wertvoll erlebt, während der Nutzen umfänglicher fachwissenschaftlichen Lehrinhalte für die spätere Berufspraxis in Frage gestellt wird.

Während die Ergebnisse der quantitativen Begleitforschung darauf hindeuten, dass ein dauerhaftes Anbieten dieses alternativen Weges in den Lehrberuf aus Sicht der Qualitätssicherung vertretbar ist, gibt die Auswertung der Interviewdaten außerdem Hinweise darauf, dass das Q-Master-Konzept auf individueller Ebene eine zeitgemäße Antwort auf zunehmend heterogene Berufsbiografien sein kann. Das Konzept scheint damit geeignet, dauerhaft und an weiteren lehrer*innenbildenden Universitäten angeboten zu werden.

Das Poster stellt aktuelle Ergebnisse der qualitativen und quantitativen der Begleitforschung zum Q-Master an der FU Berlin vor. Die Einreichenden möchten diese mit Interessierten hinsichtlich möglicher weiterführender/ergänzender Forschung und der Implikationen für die Lehrer*innenbildung diskutieren.

80. Professionelle Kompetenzen im Bereich der politischen Medienbildung bei Grundschullehramtsstudierenden – Ergebnisse einer Interventionsstudie

Mirjam Wenzel, Katharina Asen-Molz, Astrid Rank
Universität Regensburg, Deutschland

Um Medien zur Teilhabe an der Gesellschaft nutzen und mediale Inhalte kritisch beurteilen zu können, ist von jeher eine politische Medienkompetenz erforderlich (Oberle, 2017; Goll, 2018). Die Bedeutung einer politischen Medienbildung wächst in der Digitalität weiter an, da sich technologische Entwicklungen auf Gesellschaft und Politik auswirken. Die Verlagerung des Öffentlichen in digitale Sphären hat den Zugang zu Informationen und die Möglichkeiten der öffentlichen Meinungsäußerung erheblich vereinfacht (Ritzi, 2021). Gleichzeitig führt die Flut an Informationen im Internet dazu, dass deren Organisation nur noch automatisiert mittels Algorithmen bewältigbar ist und die Beurteilung der Qualität zunehmend in der Verantwortung der Konsument*innen liegt. Hinzu kommen algorithmisch personalisierte Angebote, die ein ausgewogenes Informieren erschweren und eine Polarisierung verstärken können (Schweiger, 2019; Stark et al., 2017). Nachrichten von Meinungen, Werbung oder Desinformation abgrenzen zu können sowie die Qualität von Informationen einschätzen zu können, ist eine zentrale Herausforderung in Zeiten von Fake News und Filterblasen. Da Suchmaschinen und Social Media (z.B. YouTube, TikTok) für Grundschüler*innen von enormer Bedeutung sind (mpfs, 2020) tangieren diese Phänomene zwangsläufig die Lebenswelt von Kindern, weshalb dem Sachunterricht der Grundschule die Aufgabe zukommt, entsprechende Kompetenzen anzubahnen (GDSU, 2021; Haider et al., 2022). Dazu brauchen die Grundschullehrkräfte selbst entsprechende Kenntnisse und Fähigkeiten aus verschiedenen Bezugsdisziplinen. Denn um Datafizierungs- und Algorithmisierungsprozesse reflektieren zu können oder Herausforderungen bei der Informationsbeschaffung im Internet (z. B. Desinformationen) erkennen zu können, ist informatisches, medienbezogenes und politisches Wissen erforderlich. Fachwissen allein ist für die Gestaltung qualitätsvoller Unterrichtsarrangements jedoch nicht ausreichend. Studien zu professionellen Kompetenzen von Lehrkräften haben gezeigt, dass insbesondere das fachdidaktische Wissen den Lernzuwachs der Schüler*innen beeinflusst (Kunter et al., 2010; Möller et al., 2013). Auch

Überlegungen zum Professionswissen im Sachunterricht orientieren sich an solchen kompetenzorientierten Ansätzen (GDSU, 2019). Für den Bereich der informatischen Bildung liegen bereits erste Studien vor, die ein geringes Wissen von Grundschullehrkräften und Grundschullehramtsstudierenden nahelegen (z.B. Dengel & Heuer, 2017; Gläser, 2020), zu explizit medienpolitischen Themen fehlen bislang Studien.

Das Forschungsprojekt PoliMeR (**Poli**tische **Me**dienbildung **R**egensburg) widmet sich daher der Untersuchung und Förderung des medienpolitischen Professionswissens von Grundschullehramtsstudierenden und Grundschullehrkräften. Auf dem Poster soll die Forschungsfrage beantwortet werden, ob sich das Professionswissen (Fachwissen und fachdidaktisches Wissen) von Grundschullehramtsstudierenden steigern lässt. Als Interventionsstudie im Prä-, Post-, Follow-up-Design angelegt, wird eine gemeinsame Fort- bzw. Ausbildung für Lehrkräfte und Studierende durchgeführt, in der diese kooperativ Unterrichtsideen entwickeln und in Schulklassen erproben (Asen-Molz et al., 2020). Da die Studierenden und Lehrkräfte nicht zufällig ausgewählt werden, sondern mit der freiwilligen Teilnahme an dem Seminar bzw. der Fortbildung einhergeht, handelt es sich um eine quasi-experimentelle Interventionsstudie. Zur Messung des Professionswissens sowie dessen Entwicklung durch die Intervention wird ein selbstentwickelter Online-Wissenstest mit 49 Fragen zum Fachwissen und 12 Fragen zum fachdidaktischen Wissen eingesetzt (Wenzel & Asen-Molz, i.V.). Die Pilotierungen des Instruments und der Intervention fanden im Sommersemester 2021 und Wintersemester 2021/2022 statt. Im Sommersemester 2022 wurden die erste Intervention und Haupterhebung durchgeführt und Daten im Prä- und Posttest erhoben, die Follow-up-Erhebung steht bislang noch aus. In der Experimentalgruppe nahmen 50 Personen (davon 6 Lehrkräfte) teil und 20 in der Kontrollgruppe. Zur Beantwortung der Forschungsfrage werden Mixed Models Analysen durchgeführt. Um die Stichprobe zu vergrößern, ist im Wintersemester eine zweite Haupterhebung mit Studierenden und Lehrkräften des Grundschullehramts geplant. Die bisherigen Ergebnisse der ersten Haupterhebung zeigen einen signifikanten Fachwissenszuwachs in der Experimentalgruppe, für das fachdidaktische Wissen laufen die Auswertungen aktuell noch. Das Forschungsprojekt soll erste Einblicke in das medienpolitische Professionswissen und dessen Veränderbarkeit bei Grundschullehramtsstudierenden und Grundschullehrkräften geben. Dadurch sollen spezifische Herausforderungen adressiert und zukünftig empirisch begründete Angebote zur Aus- und Fortbildung von (angehenden) Grundschullehrkräften in diesem Bereich gestaltet werden können.

81. Was wissen Lehramtsstudierende über schulischen Datenschutz? Entwicklung eines Wissenstests und erste Ergebnisse

Kim Diebig, Timo Dahlhaus, Marten Clausen
Universität Duisburg-Essen, Germany

Die Digitalisierung ist aus der heutigen Zeit nicht mehr wegzudenken. Sie hat in allen Lebensbereichen Einzug gehalten; so auch in der Schule. Nicht erst seit der COVID-19-Pandemie und damit einhergehenden Schulschließungen, in denen hybrides Lernen im Vordergrund stand, ist die Digitalisierung im schulischen Alltag angekommen. Schulische Akteure, Schülerinnen und Schüler

sowie Erziehungsberechtigte nutzen Lernplattformen und kommunizieren über digitale Medien (Denker et al., 2021).

Auch wenn der Einsatz digitaler Medien Chancen mit sich bringt, dürfen die Probleme und Herausforderungen nicht unberücksichtigt bleiben. In Bezug auf das schulische Leben stellen sich vielfältige Fragen: Darf der Vertretungsplan online abrufbar sein? Dürfen Lehrkräfte mit Erziehungsberechtigten über E-Mail kommunizieren? Dürfen Fotos von Schulausflügen auf der Schulhomepage veröffentlicht werden?

Die Fragen verdeutlichen, dass digitalisierungs- sowie datenbezogene Kompetenzen und damit einhergehende Kompetenzen im Bereich des Datenschutzes (zusammenfassend bezeichnet als Datenkompetenz) Kernanforderungen der fortschreitenden Digitalisierung sind (Denker et al., 2021; Grillenberger & Romeike, 2018), die auch für Lehrkräfte und ihr professionelles Handeln im schulischen Kontext an Bedeutung zunehmen (Beißwenger et al., 2020).

Laut Ergebnissen des Schul-Barometers, das zur Beschreibung der Schulsituation in Deutschland, Österreich und der Schweiz im März/April 2020 durchgeführt wurde, ist Datenschutz ein Hindernis dafür, dass Schulleitungen und Lehrkräfte digitale Medien und Technologien nutzen. Insbesondere hinsichtlich personenbezogener Daten liegen Unsicherheiten vor (Huber et al., 2020). 58 Prozent der befragten Lehrkräfte fühlen sich nicht ausreichend über das Thema Datenschutz informiert und 57 Prozent gaben an, dass sie über unzureichende Kompetenzen in der Anwendung digitaler Lernformate verfügen (Deutsches Schulportal, 2020).

Diese Unsicherheiten verwundern nicht, da das Thema Datenschutz in der Ausbildung von Lehrkräften bisher eine eher untergeordnete Rolle spielt. Während für die Schulbildung von bildungspolitischer Seite vielfältige Empfehlungen sowie Richtlinien formuliert werden (u.a. *Digital Competence Framework for Citizens (DigComp 2.0)*; Vuorikari et al., 2016), fehlen für die Lehrkräftebildung systematische Konzepte zur Förderung von Datenkompetenz (Grillenberger & Romeike, 2019).

Um den Blick auf die Lehrkräfteausbildung zu richten, entstand im Rahmen eines im UDE-Projekt DataCampus geförderten Lehrprojektes eine erste Version eines Wissenstests, der anschließend mit einer stärkeren Orientierung am datenschutzbezogenen Kompetenzbegriff weiterentwickelt wurde.

131 Lehramtsstudierende durchliefen den Test im Januar/Februar 2022 und beantworteten 20 Fragen im Multiple-Choice-Design zu ihrem datenschutzbezogenen Wissen. Es wird zwei Fragestellungen nachgegangen: 1. Wird das Merkmal „Datenschutz“ durch die entwickelten Wissensfragen operationalisiert (Inhaltsvalidität)? Die Inhaltsvalidität wurde von einer unabhängigen Stichprobe per Fragebogen eingeschätzt. 2. Was wissen Lehramtsstudierende über schulischen Datenschutz? Es wird davon ausgegangen, dass Lehramtsstudierende über kein fundiertes Wissen über schulischen Datenschutz verfügen. Ähnlich wie in der Lehrkräftebildung fehlen auch in der Lehrkräfteausbildung einheitliche Konzepte zur Umsetzung, sodass ein heterogenes Ergebnis erwartet wird. In Bezug auf die Inhaltsvalidität des entwickelten Testinstruments wird vermutet, dass dieses das Merkmal „Datenschutz“ breit abbilden kann. Es wird von einem mehrdimensionalen Konstrukt ausgegangen.

Mit Blick auf Fragestellung 1 ergab die Prüfung der Items in Bezug auf Inhaltsvalidität zufriedenstellende Ergebnisse, allerdings bringt die vermutete Mehrdimensionalität Herausforderungen mit sich. Bezüglich des Wissensstands Lehramtsstudierender (Fragestellung 2) lässt sich festhalten, dass diese im Durchschnitt knapp 50 Prozent der Fragen richtig beantworteten, wobei die Spanne der erreichten Ergebnisse von 15 Prozent bis 100 Prozent reicht. Während Fragen zur Kategorie „Rechtliche Grundlagen“ mit weniger als 40 Prozent richtigen Antworten bearbeitet wurden, scheint das Wissen Lehramtsstudierender in der Kategorie „Digitaler Datenschutz“ mit durchschnittlich 65 Prozent richtigen Antworten deutlich fundierter zu sein (vgl. u.a. im Spring, 2022). Die Ergebnisse zeigen Wissensbedarfe der Lehramtsstudierenden auf und helfen, Informationen zum Thema Datenkompetenz zielgenau bereitzustellen. Die Ergebnisse werden vor dem Hintergrund der Testgüte diskutiert und es werden Anregungen für weitere Forschungsarbeiten aufgezeigt.

82. Der Zusammenhang zwischen Führungsverhalten und Teamkonflikten in Kindertageseinrichtungen

Vera Eling
Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau, Deutschland

Konflikte sind unvermeidbar, wenn Menschen miteinander interagieren (Pfundmair et al., 2017). Sie treten in allen Lebensbereichen wie z.B. dem Familien- und Arbeitsleben sowie in der Gesellschaft auf (Elgoibar et al., 2017). Die Forschung zu Konflikten in sozialen Organisationen wie z.B. Kindertageseinrichtungen steckt noch in den Kinderschuhen (Simpson et al., 2017; Stövesand & Röh, 2015). Studien aus anderen Arbeitsfeldern zeigen jedoch, dass Konflikte im Team von großer Bedeutung sind. Abhängig von der Art des Konfliktes können diese sowohl nützlich (z.B. für die Kreativität und Innovation) als auch schädlich (z.B. für das Wohlbefinden und die Arbeitszufriedenheit) sein (De Dreu, 2010; De Wit et al., 2012). Auch wenn Konflikte im Team in Kindertageseinrichtungen nicht wegzudenken sind, wurden diese bisher sowohl national als auch international kaum erforscht (Eling, 2022). Unklar bleibt daher, welche Aspekte das Auftreten und Niveau von Teamkonflikten beeinflussen. Potential wird hierbei dem Führungsverhaltens der Kita-Leitung zugeschrieben (Weltzien et al., 2016; Gelfand et al., 2012). Dieses wird in der Literatur durch verschiedene positiv oder negativ konnotierte Führungsstile beschrieben. Bekannt ist hierbei, dass bei einer Führungskraft unterschiedliche Führungsstile auftreten (Einarsen et al., 2010; Kerschreiter et al., 2018).

Vor diesem Hintergrund wird in der auf dem Poster dargestellten Untersuchung der Frage nachgegangen, ob sich auf der Basis gängiger Führungsstile bestimmte Führungstypen von Kita-Leitungen ermitteln lassen. Soweit dies gegeben ist sollte untersucht werden, in welchem Zusammenhang die ermittelten Führungstypen der Kita-Leitungen mit Konflikten im Kita-Team stehen. Hierzu wurden 182 pädagogische Fachkräfte aus 75 zufällig ausgewählten Kindertageseinrichtungen aus Rheinland-Pfalz von Frühjahr bis Herbst 2020 mittels eines schriftlichen Fragebogens befragt. In diesem wurden unterschiedliche Instrumente eingesetzt, um verschiedene Arten von Teamkonflikten (Aufgaben-, Beziehungs- und Prozesskonflikte) (Lehmann-Willenbrock et al., 2011; Hüttermann, 2013) sowie die Führungsstile transformationale, transaktionale, laissez-faire/passive, authentische, katalytische und despotische Führung zu erheben (Felfe & Goihl, 2014; Franke-Bartholdt et al., 2018; Loock & Schilling, 2010). Um zu untersuchen, welche Führungstypen

unter den Kita-Leitungen vorkommen, wurde eine Clusteranalyse durchgeführt. Aufbauend auf den ermittelten Ergebnissen wurde in einer hierarchischen Regressionsanalyse der Einfluss des Führungsverhaltens der Führungskraft auf Teamkonflikte untersucht. Die Teamgröße, Dauer der Zusammenarbeit sowie die Komplexität der Betreuung wurden als Kontrollvariablen verwendet (Almost et al., 2010; Korsgaard et al., 2008).

Im Rahmen der Clusteranalyse konnten drei verschiedene Führungstypen gefunden werden. Das Cluster ‚konstruktives Führungsverhalten' vereint Kita-Leitungen, denen von ihren Mitarbeiter*innen überwiegend positiv konnotiertes Führungsverhalten zugeschrieben wurde (transformationale, authentische und katalytische Führung). Negativ konnotiertes Führungsverhalten (laissez-faire/passive und despotische Führung) zeigen die Kita-Leitungen dieses Clusters nach Einschätzung der Fachkräfte kaum. Ein ähnliches Muster zeigt sich beim Cluster ‚moderat konstruktives Führungsverhalten', wobei positiv konnotiertes Führungsverhalten in geringerem und negativ konnotiertes Führungsverhalten in geringfügig höherem Umfang gezeigt wird. Im Cluster ‚ambivalentes Führungsverhalten' wurde sowohl positiv als auch negativ konnotiertes Führungsverhalten zusammengefasst. Die Ergebnisse der Regressionsanalyse zeigen, dass konstruktive Kita-Leitungen Aufgaben-, Beziehungs- und Prozesskonflikte minimieren. Moderat konstruktive Kita-Leitungen hängen dem Gegenüber mit einem höheren Niveau an Prozesskonflikten zusammen als ambivalente Kita-Leitungen. Darüber hinaus ist der Einfluss ambivalenter Kita-Leitungen auf Beziehungskonflikte größer als der von moderat konstruktiven Kita-Leitungen.

Kita-Leitungen wird in der frühpädagogischen Diskussion eine Schlüsselfunktion für die Realisierung einer hohen pädagogischen Prozessqualität in Kindertageseinrichtungen zugeschrieben. Die Ergebnisse der Studie zeigen, dass Kita-Leitungen nach Einschätzung ihrer Mitarbeiter*innen unterschiedliche Führungsstile anwenden. Weiterhin konnte festgestellt werden, dass diese durch ihr Führungsverhalten Einfluss auf Teamkonflikte nehmen können. Hinsichtlich Konflikten im Team konnte in dieser Studie erstmals nachgewiesen werden, dass unterschiedliche Arten von Teamkonflikten in Kindertageseinrichtungen gefunden werden können. Nichtsdestotrotz verbleibt ein erhebliches Forschungsdesiderat. Zukünftige Studien sollten untersuchen, welche weiteren Faktoren das Auftreten und Niveau von Teamkonflikten beeinflussen und welche Auswirkungen diese bspw. auf die pädagogische Prozessqualität oder Zusammenarbeit im Kita-Team haben.

83. Open Innovation Mindset and School Leadership: How Innovations can be successfully implemented at schools

Jasmin Witthöft, Marcus Pietsch
Leuphana Universität Lüneburg, Deutschland

The COVID-19 pandemic caused new challenges, increased already existing problems, and underlined the importance schools play in the life of children and their individual development (United Nations, 2022). In this background, schools had to comply with governmental regulations and adapt on all levels, from innovation management over equipment to teaching methods to adequately support students' individual learning situations (Bremm et al., 2021).

However, research on innovation management revealed that schools implement innovations slowly, mainly based on firmly established and existing knowledge structures (Krieg & Haas 2021;

Goldenbaum, 2012). Especially management strategies to organize and implement targeted but flexible innovation are crucial for successful innovation (Wobser, 2022). In the business sector, the Open Innovation paradigm serves as a conceptual approach to design innovation processes. Open Innovation is "a distributed innovation process based on purposively managed knowledge flows across organizational boundaries, using pecuniary and non-pecuniary mechanisms in line with the organizations' business model" (Chesbrough & Bogers, 2014, p. 17). Organizations must become more transparent, open their innovation processes, collaborate with external partners and knowledge sources, and see beyond internal standards and resources (Mergel, 2018). In the dynamic and interdependent structures of schools, school leaders are seen as drivers of innovation (Bonsen, 2016) and boundary spanners who foster knowledge exchange across school boundaries (Benoliel & Schechter, 2017). Generally, they are responsible for balancing the realization of educational standards and supporting students' individual development. Investigations on leadership revealed that leaders' convictions, motives, and mindsets could determinately impact decision-making processes, teaching and learning techniques, and the awareness of cooperation and innovation within an organization (Engelsberger et al., 2022; King & Anderson, 2002; Osborn & Brown, 2005). According to Engelsberger et al. (2022), Open Innovation can be enhanced through leaders' Open Innovation Mindset, a multidimensional construct of dynamic capabilities, i.e., openness, creativity, attitudes towards knowledge sharing, as well as risk and failure tolerance. Whereas Open Innovation and dynamic managerial capabilities are well established in the economic discourse implemented by several companies, it has not been extensively investigated in the educational discourse (Bigliardi et al., 2020).

Against this background, the research question: *Does the Open Innovation mindset of school leaders affect the innovation performance at German schools? If so, how?,* guided this investigation. For this purpose, data from the third wave of the Leadership in Germany Schools (LineS) study, a principal panel (N=411) that is representative of Germany (Cramer et al., 2021), is re-analyzed. First, a factor analysis showed that the multidimensional Open Innovation Mindset, which was previously conducted among managers in commercial enterprises (Engelsberger et al., 2022), can be demonstrated among German school leaders. The Open Innovation Mindset encompasses the following dynamic capabilities: openness, creativity, positive attitude toward knowledge sharing and sensing, and risk and failure tolerance.

The investigation revealed that the Open Innovation Mindset of school leaders and the innovation performance, which was measured by the fact that rather radical innovations were introduced at the classroom level of a school in the 12 months prior to the survey, at German schools are positively correlated (r=.12, p<.01). Third, it is striking that no direct significant correlation of the Open Innovation Mindset of school leaders with the open innovation breadth (i.e., the number of external sources of knowledge, e.g., universities, authorities, other schools etc.) and the open innovation depth (i.e., the total amount of knowledge that comes from external sources) of a school can be proven (p> .05). Consequently, school leaders predominantly draw their knowledge to innovate from internal sources. External knowledge sources are not used extensively. The results underline the limited exchange between schools and external partners.

(595 of 600 words)

84. Die wahrgenommene Verantwortung von Lehrenden in der berufsbezogenen Weiterbildung für kurzfristige und langfristige Bildungsergebnisse

Susanne Wißhak, Dorothee Barth
Universität Konstanz, Germany

Theoretischer Hintergrund

Wenn sich Personen für ein Ergebnis verantwortlich fühlen, steigert dies ihre Motivation (Hackman & Oldham, 1976), ihr Engagement und Durchhaltevermögen (Halvorsen et al., 2009) wie auch ihre Performanz (Hackman & Oldham, 1976; Humphrey et al. 2007) und ihre berufliche Zufriedenheit (Matteucci et al. 2017; Winter at al., 2006).

Wie für schulischen Unterricht, gilt auch für berufsbezogene Weiterbildungsmaßnahmen, dass die Lehrenden Einfluss auf die Lehr-Lernqualität nehmen und somit auch die Ergebnisse von Bildungsmaßnahmen beeinflussen können (Burke & Hutchins, 2008; Grohmann et al., 2021). Im Hinblick auf Unterricht betonen Lauermann und Karabenick (2013), dass dessen Erfolg nicht nur von dem Wissen und Können der Lehrenden abhängt, sondern auch von ihrer empfundenen Verantwortung. Dies entspricht auch gängigen Kompetenzkonzepten und -modellen für Lehrkräfte, in welchen neben Fähigkeiten auch Überzeugungen sowie motivationale und volitionale Bereitschaften für wichtig erachtet werden (Baumert & Kunter, 2006; Weinert, 2002). Hinsichtlich der Verantwortung für guten Unterricht weisen Lauermann und Karabenick (2011) zudem darauf hin, dass zwischen der persönlich erlebten und der von außen zugewiesenen Verantwortung unterschieden werden muss.

Für die berufsbezogene Weiterbildung, in welcher erwachsene Lernende in sogenannten Trainings beruflich relevante Kompetenzen erwerben sollen, wird zwischen vier Ebenen von Bildungsergebnissen unterschieden: (1) Reaktionen (z.B. Zufriedenheit), (2) Lernerfolg, (3) Transfer (i.S. einer Anwendung am Arbeitsplatz) sowie (4) Resultate auf der Ebene der Organisationen, bei welchen die Lernenden beschäftigt sind (Kirkpatrick & Kirkpatrick, 2012). Die persönliche wie auch die zugewiesene Verantwortung der Lehrenden für diese Weiterbildungsergebnisse waren bisher noch nicht untersucht.

Fragestellung

In dem Beitrag wird der Frage nachgegangen, in welchem Maß Lehrende in der berufsbezogenen Weiterbildung für verschiedene Weiterbildungsergebnisse *persönliche* Verantwortung wahrnehmen und *von außen zugewiesene* Verantwortung verspüren. Zudem stellt sich die Frage nach der Dimensionalen Struktur der beiden Konstrukte.

Methode

Für die Studie wurden mittels eines Online-Fragebogens $N = 393$ Lehrende in der berufsbezogenen Weiterbildung dazu befragt, wie sie ihre persönliche und zugewiesene Verantwortung für Weiterbildungsergebnisse bewerten. Aufgrund fehlender Instrumente wurden, basierend auf den o.g. Ebenen nach Kirkpatrick und Kirkpatrick (2012), für beide Konstrukte je acht Items entwickelt. Die Items wurden dabei für beide Konstrukte identisch formuliert, nur der Itemstamm variierte zwischen der persönlichen und der zugewiesenen Verantwortung, z.B.:

Itemstamm für die persönliche Verantwortung: „*In meinem letzten Training fühlte ich mich als Trainer/in persönlich verantwortlich dafür, ...*“

Item 6: „*...dass die Teilnehmenden das neu Erlernte erfolgreich bei ihrer Arbeit umsetzen*“.

Zur Überprüfung der Dimensionalität wurden zunächst für beide Konstrukte explorative Faktorenanalysen mit Varimaxrotation vorgenommen. Für die identifizierten Faktoren wurden Mittelwerte gebildet, mithilfe von t-Tests wurden die Ausprägungen der einzelnen Dimensionen der Verantwortung verglichen.

Ergebnisse und ihre Bedeutung

Sowohl die persönliche (M = 8.05) als auch die zugewiesene Verantwortung (M = 8.35) wurden von den Lehrenden auf der Skala von 0 bis 10 als hoch eingeschätzt. Die explorative Faktorenanalyse ergab für beide Konstrukte jeweils zwei Dimensionen, die *Verantwortung für kurzfristige Trainingsergebnisse* (Reaktionen und Lernerfolg) sowie die *Verantwortung für langfristige Trainingsergebnisse* (Transfer und organisationale Resultate). Mittelwertvergleiche für die *persönliche* Verantwortung zeigten, dass die Lehrenden hinsichtlich kurzfristiger Ergebnisse (M = 8.95; SD = 1.18) mehr Verantwortung verspüren als für langfristige Ergebnisse (M = 7.16; SD = 2.43; p < .001; d = 0.82). Für die *zugewiesene* Verantwortung ergab sich der gleiche Unterschied, wenn auch mit einer geringeren Effektstärke (p < .001; d = 0.42). Ferner zeigte sich, dass bezüglich langfristiger Ergebnisse die zugewiesene Verantwortung höher ausfällt als die persönliche Verantwortung (p < .001, d = 0.33).

Die Befunde legen nahe, dass vor allem die persönliche Verantwortung Lehrender für langfristige Weiterbildungsergebnisse gestärkt werden sollte, damit Weiterbildungsanstrengungen zukünftig nachhaltigere Wirkungen erzielen.

85. Individuelle und kontextuelle Prädiktoren für transferrelevante Einstellungen in Professionalisierungsprozessen

Mareike Ehlert, Elmar Souvignier
WWU Münster, Germany

Theoretischer Hintergrund

Für den Transfer evidenzbasierter Interventionen in den Unterricht stellen Professionalisierungsmaßnahmen wie Fortbildungen für Lehrkräfte eine wichtige Unterstützung dar (Desimone, 2009). Gewünschte Effekte von Professionalisierungsmaßnahmen umfassen dabei – neben der Vermittlung fachlicher Kompetenzen – die Entwicklung positiver Einstellungen gegenüber der Intervention wie z.B. die Akzeptanz und die wahrgenommene Machbarkeit (Petermann, 2014). Solche positiven Einstellungen gehen mit einer erhöhten Nutzung (Hebbecker et al., 2020) und mit verbesserten Schüler:innenleistungen einher (Henninger, 2010) und bilden sich relativ früh im Transferprozess heraus (Petermann, 2014). Für den Transfererfolg ist ebenso die intendierte Wiedergabetreue entscheidend, d.h. die Absicht einer Lehrkraft, die Intervention wie vorgesehen im Unterricht umzusetzen (Petermann, 2014). In Professionalisierungsmaßnahmen sollte ein hohes Maß an Bereitschaft für eine wiedergabetreue Umsetzung der Konzepte erzielt werden, da dies die Wirksamkeit der Intervention selbst beeinflusst (Kuhn, 2018).

Allerdings zeigen sich in Befragungen zu transferrelevanten Einstellungen häufig hohe interindividuelle Unterschiede zwischen Lehrkräften (Hebbecker et al., 2020). Diese könnten sich zum Teil durch kontextuelle Merkmale erklären lassen: Auf Schulebene sind z.B. die Kooperation zwischen Lehrkräften (Postholm, 2012), die wahrgenommene Unterstützung durch die Schulleitung (Banilower et al., 2007) sowie das Transferklima (Grossman & Salas, 2011) relevante Prädiktoren für den Kompetenzaufbau von Lehrkräften und für die Anwendung innovativer Praktiken. Deren Einfluss auf Akzeptanz, Machbarkeit und Wiedergabetreue wurde bislang allerdings noch nicht geprüft. Neben den kontextuellen können auch individuelle Merkmale der Lehrkräfte transferrelevante Einstellungen beeinflussen. Wie stark eine Lehrkraft von einer Innovation überzeugt ist, ist z.B. abhängig von ihrer Selbstwirksamkeitserwartung (Donnell & Gettinger, 2015), kongruenten Einstellungen zum Lerninhalt sowie von ihrem selbstberichteten Wissen über die Intervention (Lakin & Shannon, 2015). Wenig bekannt ist hingegen über den Einfluss dieser Merkmale auf weitere transferrelevante Einstellungen wie Machbarkeit und Wiedergabetreue. Ziel dieser Studie ist deswegen, den Einfluss kontextueller und individueller Merkmale auf Akzeptanz, Machbarkeit und Wiedergabetreue zu untersuchen.

Fragestellungen

1. Inwiefern beeinflussen kontextuelle Rahmenbedingungen (Schulleitung, Kooperation, Transferklima) transferrelevante Einstellungen von Lehrkräften in Professionalisierungsprozessen (Akzeptanz, Machbarkeit, Wiedergabetreue)?

H1: Lehrkräfte mit hoher Unterstützung durch die Schulleitung, guter kollegialer Kooperation und positivem Transferklima zeigen eine höhere Akzeptanz, Machbarkeit und intendierte Wiedergabetreue gegenüber der Intervention.

2. Inwiefern beeinflussen individuelle Voraussetzungen (Selbstwirksamkeit, kongruente Einstellung zum Lerninhalt, selbstberichtetes Wissen) transferrelevante Einstellungen in Professionalisierungsprozessen (Akzeptanz, Machbarkeit, Wiedergabetreue)?

H2: Lehrkräfte mit hoher Selbstwirksamkeitserwartung, kongruenten Einstellungen und hohem selbstberichteten Wissen über den Lerngegenstand zeigen eine höhere Akzeptanz, Machbarkeit und intendierte Wiedergabetreue gegenüber der Intervention.

Methode und Ergebnisse

Zunächst wurden *N*=12 Lehrkräfte über einen Zeitraum von sechs Monaten als Fortbildende qualifiziert. Seit September bilden diese durch drei ganztägige Präsenzveranstaltungen sowie durch die eigenständige Bearbeitung eines Blended-Learning-Kurses *N*=68 Grundschullehrkräfte an 27 Schulen fort. Diese bilden die Stichprobe dieser Untersuchung. Die teilnehmenden Lehrkräfte sind größtenteils weiblich (83%), im Mittel 38.56 Jahre alt (*SD*=9.14) und seit *M*=9.89 Jahren als Lehrkraft tätig (*SD*=8.81). Gegenstand der Fortbildung ist eine evidenzbasierte Intervention zur differenzierten Leseförderung im Grundschulunterricht, für die ausgearbeitete Unterrichtsmaterialien bereitgestellt werden. Die Blended-Learning-Einheiten zur Fortbildung der Lehrkräfte bieten multimodal aufbereitete theoretische Inputs, beispielsweise in Form von Erklärvideos. Zudem werden aktive Lerngelegenheiten durch Selbsttests und Portfolioaufgaben geschaffen. Dadurch sind die Lehrkräfte bei Fortbildungsabschluss mit der Leseintervention gut vertraut und können Einschätzungen zu

transferrelevanten Einstellungen geben, haben die Intervention aber noch nicht im Unterricht durchgeführt.

Bisher liegen die Lehrkraft-Daten des ersten Messzeitpunkts vor Fortbildungsbeginn vor, bei dem individuelle und kontextuelle Prädiktoren erfasst wurden. Die Skalen zeigen hohe interne Konsistenzen auf (α=.79-.93). Alle Konstrukte wurden mittels standardisierter Instrumente erhoben. Im Dezember folgt nach Abschluss der Fortbildung die zweite Erhebung zu transferrelevanten Einstellungen. Der Einfluss der Prädiktoren soll dann regressionsanalytisch ausgewertet werden. Die Ergebnisse geben Aufschluss darüber, was transferrelevante Einstellungen hinsichtlich individueller oder kontextueller Faktoren charakterisiert. Dies liefert wichtige Hinweise auf passgenaue Unterstützungsmaßnahmen, die die Komplexität von Transferprozessen berücksichtigen.

86. Guter Unterricht durch kooperatives Lernen: Eine Evaluationsstudie zur Förderung des kooperativen Lernens an der Gemeinschaftsschule

Johanna Marder[1], Alexandra Dehmel[1], Katja Adl-Amini[2], Vanessa A. Völlinger[3], Benjamin Fauth[1]
[1]Institut für Bildungsanalysen Baden-Württemberg; [2]Technische Universität Darmstadt; [3]Justus-Liebig-Universität Gießen

Theoretischer Hintergrund

Kooperatives Lernen ist ein vielversprechendes pädagogisches Konzept, um sowohl fachliche als auch soziale Kompetenzen der Schülerinnen und Schüler (SuS) zu fördern (Slavin et al., 2003; Kyndt et al., 2013). Durch die gemeinsame Zusammenarbeit der SuS kann die Lehrkraft gezielt lernförderliche Interaktionen zwischen den SuS organisieren und unterstützen (Webb, 2009). Darüber hinaus bietet kooperatives Lernen ein großes Potential zum Umgang mit heterogenen Lernvoraussetzungen der SuS (Büttner et al., 2012), was einen großen Mehrwert für den Unterricht an der Gemeinschaftsschule als integrative Schulform mit großer Heterogenität verspricht.

Auch aus theoretischer Sicht bietet kooperatives Lernen zahlreiche Anknüpfungspunkte, einen kognitiv aktivierenden, konstruktiv unterstützenden Unterricht zu realisieren und so zu einem qualitätsvollen Unterricht im Sinne der Tiefenstrukturen (Klassenführung, konstruktive Unterstützung, kognitive Aktivierung; Hamre & Pianta, 2001; Klieme et al., 2009) beizutragen.

Viele Lehrkräfte sehen sich jedoch mit Herausforderungen bei der Anwendung des kooperativen Lernens in der Unterrichtspraxis konfrontiert, wie z.B. eine erschwerte Klassenführung oder unzureichende soziale Kompetenzen der SuS (Adl- Amini et al., 2022) und setzen kooperatives Lernen vergleichsweise wenig ein (Abramczyk & Jurkowski, 2020; Völlinger et al., 2018).

Diese Befunde sind für die Bildungsforschung in mehrfacher Hinsicht relevant: Sie verdeutlichen den Bedarf nach einer differenzierten empirischen Untersuchung, in wie weit Basiselemente des kooperativen Lernens mit den Tiefenstrukturen des Unterrichts zusammenhängen. Darüber hinaus werfen sie die Frage auf, wie insbesondere Lehrkräfte der Gemeinschaftsschulen darin unterstützt werden können, kooperatives Lernen gewinnbringend langfristig einzusetzen und so einen qualitätsvollen Unterricht zu realisieren. Einen zentralen Ansatzpunkt stellt dabei die Entwicklung der Lehrkraftkompetenz dar (Völlinger et al., 2018).

Fragestellung

Das Poster stellt eine laufende Evaluationsstudie einer Fortbildung für Lehrkräfte („*Kooperatives Lernen unter Berücksichtigung der Tiefenstrukturen an Gemeinschaftsschulen*“) in Baden-Württemberg vor. Die Fortbildung besteht aus mehreren Modulen zum kooperativen Lernen, welche über einen Zeitraum von ca. 5 Monaten in einem Blended-learning Format stattfinden. Dadurch werden Fortbildungseinheiten und Unterrichtspraxis miteinander verknüpft und unterstützt den Transfer der Fortbildungsinhalte in den alltäglichen Unterricht.

Erste Ergebnisse zu folgenden Forschungsfragen werden präsentiert:

1) Welche Bedingungen in der Gemeinschaftsschule erleben Lehrkräfte als förderlich, bzw. hinderlich bei der Umsetzung des kooperativen Lernens im Unterricht?

2) In wie weit tragen die Basiselemente des kooperativen Lernens zu einem qualitätsvollen Unterricht im Sinne der Tiefenstrukturen bei?

3) Treten Veränderungen beim Wissen, Überzeugung, sowie dem quantitativen und qualitativen Einsatz des kooperativen Lernens im Unterricht während des Fortbildungsprozesses auf?

Methode

Seit 03/2022 haben Lehrkräfte der Gemeinschaftsschulen in Baden-Württemberg die Möglichkeit, an einer Fortbildungsreihe zum kooperativen Lernen teilzunehmen. Lehrkräfte der dadurch gewonnenen Stichprobe (ca. N = 100 Lehrkräfte) unterscheiden sich hinsichtlich Berufserfahrung, Unterrichtsfächer und Vorerfahrung beim kooperativen Lernen. Da diese Merkmale als relevant für die Umsetzung des kooperativen Lernens gelten (Völlinger et al., 2018), scheint die Stichprobe zur Untersuchung der Fragestellungen geeignet.

Die Evaluation ist in einem längsschnittlichen Design angelegt und umfasst den gesamten Fortbildungsprozess sowie ein zusätzliches Follow-up nach Abschluss der Fortbildung. Mit Hilfe von Online-Fragebögen werden Lehrkräfte an insgesamt vier Messzeitpunkten zu ihrem (Vor-) Wissen, Überzeugungen, sowie zur Umsetzungshäufigkeit und Umsetzungsqualität des kooperativen Lernens in Ihrem Unterricht befragt. Darüber hinaus werden die Lehrkräfte gebeten, ihren Unterricht anhand der Dimensionen Klassenführung, konstruktive Unterstützung und kognitive Aktivierung einzuschätzen. Zur Beantwortung der Fragestellungen kommen deskriptive, korrelations- und regressionsanalytische Auswertungsmethoden zum Einsatz.

Ertrag der Evaluationsstudie

Die gewonnenen Erkenntnisse ermöglichen Einblicke in Gelingensbedingungen und Umsetzungsschwierigkeiten des kooperativen Lernens in der Gemeinschaftsschule. Die empirische Untersuchung des Zusammenhangs der Basiselemente kooperativen Lernens mit Tiefenstrukturen soll ein differenziertes Bild ermöglichen, in wie weit kooperatives Lernen zu einem qualitätsvollen Unterricht beiträgt und ermöglicht die Identifikation von Stellschrauben zur Unterrichtsentwicklung. Durch die Ergebnisse sollen konkrete Unterstützungsmaßnahmen für Lehrkräfte der Gemeinschaftsschule abgeleitet werden und in die Weiterentwicklung der Fortbildungsmaßnahme einfließen.

87. Selbstwirksamkeit und Teamfähigkeit in kooperativen Unterrichtsphasen in der inklusiven Grundschule fördern: Ist die SEGEL-Intervention der Schlüssel?

Stefanie Bosse, Nadine Spörer
Universität Potsdam, Germany

Das Sozialverhalten von Lernenden ist beeinflusst durch soziale und personale Kompetenzen sowie soziale Einstellungen (Kunter & Stanat, 2002). Besonders in der inklusiven Grundschule, in der Kinder mit teilweise sehr heterogenen Voraussetzungen lernen, kann das Sozialverhalten Einzelner zur Belastung für die Klassengemeinschaft und das Klassenklima werden (Knickenberg, Stöcker & Zurbriggen, 2021). Insofern erscheint es notwendig, spezifische Kompetenzfacetten unterrichtsintegriert zu fördern. Ein möglicher Ansatz können kooperative Lernformen sein (Syring, 2019), die einerseits die Teamfähigkeit (als eine Facette der Sozialkompetenzen) fokussieren, jedoch auch eine gruppenspezifische Selbstwirksamkeit (als ein Aspekt der personalen Kompetenzen) fördern.

Der Selbstwirksamkeit wird vor allem im Zuge der Professionalisierung von Erwachsenen eine hohe Bedeutsamkeit zugesprochen. So wirken zum Beispiel Erfolgserlebnisse und positive Vorbilder als Quellen der Selbstwirksamkeit (Schwarzer & Jerusalem, 2002). Bei Lernenden werden ähnliche Mechanismen angenommen, empirisch überprüft werden sie jedoch selten. Ähnlich verhält es sich im Bereich der Teamfähigkeit, die bei Erwachsenen empirisch untermauert ist, bei Kindern und Jugendlichen hingegen kaum. Dieses Desiderat greift die Interventionsstudie „Soziale Entwicklung beim gemeinsamen Lernen“ (SEGEL) auf.

Folgende Fragestellung lenkt die Studie: Inwieweit können die sozialen und personalen Kompetenzen und die Einstellungen von Grundschulkindern zum gemeinsamen Lernen mit Hilfe einer unterrichtsintegrierten Intervention verbessert werden? Im Hinblick auf den hier vorliegenden Beitrag wird die Fragestellung spezifiziert: Wie entwickeln sich die Teamfähigkeit und die gruppenspezifische Selbstwirksamkeit infolge der SEGEL-Intervention?

Zur Beantwortung der Fragestellung wurde eine manualbasierte Kurzintervention (6 x 90 Minuten) für 4. und 5. Jahrgangsstufen unterrichtsintegriert durchgeführt. Es kam ein Switching-Replications-Design mit zwei Gruppen bzw. parallelen Klassen zum Einsatz (Shadish, Cook & Campbell, 2002). Die erste Gruppe erhielt die Intervention durch vorab geschulte Trainer:innen, während die zweite Gruppe regulär durch die Klassenlehrkraft unterrichtet wurde. Nach zwei Wochen erfolgte ein Wechsel. Jede Interventionssitzung adressierte vorab ausgewählte Kompetenzen (Bosse & Spörer, 2022) und sollte Lernenden dabei unterstützen, sich in andere Lernende und in sich selbst einzufühlen, kooperativ zu arbeiten und die Klassengemeinschaft zu wertschätzen.

Der Interventionserfolg wurde zu drei Messezeitpunkten (t_0, t_1, t_2) mittels Fragebögen erfasst. In den Fragebögen wurden unter anderem die Teamfähigkeit (3 Items; Furthenmüller, 2015) und die Gruppen-Selbstwirksamkeit (4 Items; Beierlein, Kovaleva, Kemper & Rammstedt, 2014) auf einer vierstufigen Skala (von 0 = *stimmt gar nicht* bis 3 = *stimmt genau*) erfragt. Darüber hinaus schätzten die Lernenden nach jeder der sechs Interventionssitzung ihre Selbstwirksamkeit bezogen auf die Klasse (2 Items, Eigenentwicklung) auf einer 10 cm langen stufenlosen Linie ein.

Bis zur Einsendung des Abstracts lagen Daten von sieben Klassen (Gruppe 1 = 90 Lernende, Gruppe 2 = 63 Lernende) vor. Während sich in Bezug auf die Teamfähigkeit keine Pretestunterschiede zwischen den Gruppen zeigten, fanden sich diese bei der Gruppen-Selbstwirksamkeit. Es lagen teilweise hohe Ausgangswerte vor: Gruppen-Selbstwirksamkeit M_{G1} = 2.10 (*SD* = .54) bzw. M_{G2} = 1.86 (*SD* = .64), Teamfähigkeit: M_{G1} = 2.09 (*SD* = .61), M_{G2} = 1.95 (*SD* = .58). Infolge der Intervention verbesserten sich die Werte der Gruppen-Selbstwirksamkeit in der Gruppe 1 signifikant, alle anderen Werte blieben stabil. Die Veränderung über die Zeit in Bezug auf das Maß der stufenlosen Linie (Selbstwirksamkeit bezogen auf die Klasse) wurde mit Hilfe einer Varianzanalyse mit Messwiederholung überprüft. Es zeigte sich keine signifikante Veränderung.

Insgesamt ergeben sich bislang kaum Nachweise in Bezug auf die Wirksamkeit der Förderung. Dieses Ergebnis sollte dennoch Erwähnung finden und diskutiert werden. Bis zur Tagung werden Daten von voraussichtlich weiteren 8 Klassen vorliegen. Es wird sich dann zeigen, ob die Ergebnisse der größeren Stichprobe denen der bisherigen Datenbasis entsprechen.

88. Social Media Influencing: Entwicklung einer Intervention zur Förderung von Influencer*innen-Resilienz bei Jugendlichen

Josephine Jahn[1], Tamara Marksteiner[1], Vogel Tobias[2], Heinz Aaron[2], Duffner-Korbee Dorien[3], Rüther Lale[1], Henninger Carolin[2]

[1]Universität Mannheim, Germany; [2]Hochschule Darmstadt; [3]Fraunhofer-Institut für System- und Innovationforschung ISI

Theoretischer Hintergrund

Die Informationsfülle in den Sozialen Medien kann überfordern, da Relevantes von Irrelevantem nicht mehr unterschieden werden kann. Hier kommt Influencer*innen eine besondere Rolle zu: sie nutzen die Sozialen Medien zur (bezahlten) Bewerbung von Produkten und Marken und reichern so das Geflecht aus gesichertem Wissen und Meinung an.

Bisherige Forschungsergebnisse legen nahe, dass sich dieses Influencer*innen-Marketing negativ auf das Kaufverhalten sowie das Wohlbefinden von Jugendlichen auswirkt.

Zudem weisen bestehende Befunde darauf hin, dass manche Jugendliche gegenüber der Risikosituation des Influencer*innen-Marketings resilienter sind als andere. Sie legen nahe, dass Jugendliche mit hohem Selbstwert und hoher Advertising Literacy weniger anfällig für die Beeinflussung durch Influencer*innen-Marketing sind. Zudem konnten Analysen feststellen, dass Resilienz bei verschiedensten Risikosituationen eine abmildernde Wirkung hat.

Die Förderung der Faktoren, welche die negativen Folgen von Influencer*innen-Marketing abmildern, kann in der pädagogischen Praxis durch die Durchführung einer wissenschaftlich fundierten Intervention erfolgen. Bislang gibt es eine derartige Intervention jedoch noch nicht. Empirische Erkenntnisse zu den Moderationseffekten auf die Auswirkungen von Influencer*innen-Marketing liefern Hinweise für die Entwicklung einer effektiven Intervention.

Fragestellung

Daher wird in einem vom Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz geförderten Verbundprojekt in einem ersten Schritt erforscht, welche Faktoren die

negativen Konsequenzen von Influencer*innen-Marketing abmildern und somit als Schutzfaktoren, im Sinne einer Influencer*innen-Resilienz, gelten.

In einem zweiten Schritt wird darauf aufbauend, ein pädagogisch-psychologisches Interventionsmanual entwickelt, dass von (Sozial-)Pädagog*innen zur Förderung von Influencer*innen-Resilienz verwendet werden kann.

Somit werden im Verbundprojekt die folgenden Fragestellung untersucht:

*1) Wie stark moderieren Selbstwert, Advertising Literacy und Generelle Resilienz bei Jugendlichen den Zusammenhang zwischen dem Konsum von Influencer*innen-Marketing und dem maladaptiven Konsumverhalten sowie dem Wohlbefinden?*

Angenommen wird, dass Selbstwert, Advertising Literacy und Generelle Resilienz (1) die Stärke des positiven Zusammenhangs von Influencer*innen-Marketing und maladaptivem Konsumverhalten und (2) die Stärke des negativen Zusammenhangs von Influencer*innen-Marketing und Wohlbefinden abschwächt.

*2) Wie sieht eine empirisch fundierte Intervention aus, die der Förderung von Influencer*innen-Resilienz dient?*

Angenommen wird, dass die gezielte Förderung von Selbstwert, Advertising Literacy und Genereller Resilienz im Rahmen einer Intervention die negativen Folgen von Influencer*innen-Marketing abmildert.

Methode

Zur Beantwortung der ersten Fragestellung werden insgesamt 1300 Jugendliche im Alter von 14 bis 18 Jahren mithilfe eines Online-Fragebogen befragt.

Zur Messung des Konsums von Influencer*innen-Marketing wird den Teilnehmenden ein Online-Fragebogen mit Skalen u.a. zum Konsum der Beiträge von Influencer*innen und zur wahrgenommene Produktplatzierung vorgelegt. Das psychologische Konstrukt Wohlbefinden wird durch Skalen zur Messung der Lebenszufriedenheit und Depressivität erfasst. Zur Abbildung des maladaptiven Konsumverhaltens wurde eine neue Skale entwickelt, wohingegen die untersuchten Moderatorvariablen mithilfe etablierter Skalen erhoben werden.

Die Testung der Moderationshypothesen erfolgt regressionsanalytisch und wird bis zur Tagung abgeschlossen sein, sodass die Ergebnisse dann vorgestellt werden können.

Zur Beantwortung der zweiten Fragestellung werden die Befunde zur ersten Fragestellung interpretiert und in die Entwicklung von Interventionsinhalten überführt. In der Intervention sollen die Merkmale der Jugendlichen besonders gefördert werden, die in der Analyse zur ersten Fragestellung eine abmildernde Wirkung des negativen Einflusses des Influencer*innen-Marketing gezeigt haben.

Ergebnisse und ihre Bedeutung

Die identifizierten Moderatoren können Aufschluss darüber geben, welche Merkmale von Jugendlichen das Risiko für negative Folgen des Influencer*innen-Marketings abmildern. Mithilfe dieser identifizierten Schutzfaktoren werden Interventionsansätze zur Förderung von Resilienz bei

Jugendlichen im Bereich des Influencer*innen-Marketings abgeleitet. Diese werden in einem Manual aufbereitet, das (Sozial-)Pädagog*innen zur Verfügung gestellt wird. Einen ersten Entwurf des Manuals stellen wir auf der Tagung zur Diskussion.

Somit tragen die Forschungsergebnisse dazu bei, dass Fachkräfte eine wissenschaftlich fundierte Möglichkeit erhalten, die Jugendlichen im Geflecht aus gesichertem Wissen und Gerücht, das Soziale Medien bereitstellen, zu begleiten. Zudem leistet das Forschungsvorhaben einen Beitrag dazu, die Wirkungsorientierung von praxisrelevanten Interventionen in der pädagogischen Arbeit weiter zu fördern.

89. Outdoor-Training - Selbstkonzept und Umgang mit Stress und Belastung

Yannic Ebner, Thomas Eberle
Friedrich-Alexander-Universität Erlange-Nürnberg, Germany

Nicht nur der Krieg in der Ukraine auch die Folgen der Corona Pandemie schärfen den Blick für globale Lieferketten und stießen in Deutschland eine Digitalisierungsoffensive an. Gleichzeitig wächst auf Grund von Rohstoffengpässen und Naturkatastrophen (z.B. Hochwasser Ahrtal, 2022) das Bedürfnis nach mehr Sensibilität im Umgang mit den natürlichen Ressourcen der Erde. Schnelle Veränderungen erfordern verantwortungsbewusstes Handeln unter Stress. Handlungsorientiertes Lernen kann im Rahmen erlebnispädagogischer Lehr-Lern-Szenarien zum Training dieser Fähigkeit beitragen.

Rousseau benennt in seiner Erziehung über den Menschen drei zentrale Aspekte: „Die Natur oder die Dinge oder die Menschen“ (Rousseau, 1975, S.10). Für ihn ist das Erlebnis Ausgangspunkt gesteigerter Selbstwahrnehmung und Begegnungspunkt zwischen Subjekt und Persönlichkeit (Fischer & Ziegenspeck, 2000, S. 104). Vor dem Hintergrund gesellschaftlicher Individualisierungspraktiken plädiert Newman (2016) für eine veränderte Sichtweise auf das Subjekt. Er beschreibt dieses als *Singularität* (Newman, 2016) und versteht es im Sinne von Stirner (1986) als etwas Unbestimmbares bzw. nicht Repräsentierbares. Newman (2016) leitet daraus eine Unbeherrschbarkeit als Grundprämisse des Einzelnen ab und plädiert in seinen Werken für mehr Autonomieförderung. Unbekannte Situationen erfordern autonome Entscheidungen, weshalb sich unter professioneller Führung ermöglichte Grenzerfahrungen (Eberle, 2019) in erlebnispädagogischen Settings zu Trainingszwecken dieser Art besonders eignen. Vor allem Natursportarten (z.B. Kajak, Mountainbiking) bieten ideale Vorrausetzungen, da sie die Auseinandersetzung mit individuellen Verhaltensmustern anregen, von externen Faktoren (z.B. Wetterlage, Ausrüstung) beeinflusst werden und Situationen unter Stressbedingungen erzeugen können. In den Trainings werden deshalb Phasen der Aktion und Reflexion integriert, um einen Raum für autonome Persönlichkeitsbildung zu schaffen. Der Umgang mit Stress und Belastung wird dabei zum zentralen Reflexionsparameter.

In dem Dissertationsprojekt „Outdoor-Training – Selbstkonzept und Umgang mit Stress & Belastung“ wird deshalb untersucht, inwiefern sich Aspekte des Selbstkonzepts über das Training verändern und wie die Teilnehmer*innen mit Stress und Belastung während des Trainings umgehen. Das Vorhaben knüpft an bereits bestehenden Forschungsarbeiten (z.B. Eberle, 2004a; Markus et. al, 2016a) an. Empirische Untersuchungen berichten immer wieder von positiven Selbstkonzeptveränderungen (z.B.

Cook, 1980; Moch, 2002; Eberle, 2002; Boeger, 2005a; Tuch, 2017). Ähnliche Ergebnisse lassen sich in der Literatur zahlreich widerfinden (Dutkiewicz & Chase, 1991; Davis-Berman et al, 1994; Wheeler, Goldie & Hicks, 1998). Im Hinblick auf den Umgang mit Stress und Belastung zeigt sich jedoch dringender Forschungsbedarf.

Aus diesem Grund wurden *N* = 208 Teilnehmer*innen zu Aspekten des Selbstkonzepts (Selbstsicherheit im Alltag, Selbstwirksamkeit) sowie zum Umgang mit Stress und Belastung befragt. Die im Rahmen hochschuldidaktischer Seminare durchgeführten Outdoor-Trainings umfassen klassische Natursportarten wie Kajak oder Mountainbiking und wurden vorrangig mit Lehramtsstudierenden (Grund,- Mittel,- und Realschule) der Friedrich-Alexander-Universität Erlangen-Nürnberg durchgeführt. Kurz- und Langzeittrainings wurden gleichermaßen berücksichtigt.

Für die Selbstkonzeptskalen wurde ein Pre-Post-Design (Fragebogen) gewählt, während für die offenen Fragen zum Umgang mit Stress ein Post-Test Verfahren (offene Fragen - Fragebogen) eingesetzt wurde. Die erhobenen Daten wurden nach der Erhebung in die Software IBM SPSS Statistics übertragen. In den Fragebögen sollten sich die Teilnehmer*innen bzgl. Selbstsicherheit im Alltag (Crone & Schimmelpfennig, 1983) und allgemeiner Selbstwirksamkeit (Jerusalem & Schwarzer, 1986) auf einer fünfstufigen Ratingskala (Likert-Skala, vgl. Rost, 1996) selbst einschätzen. Über die zwei Messzeitpunkte wird eine positive Veränderung des Selbstkonzepts erwartet, dass mit Hilfe von Mittelwertvergleichen (t-Test) geprüft wird. Zur Klärung der verschiedenen Bewältigungsstrategien wurden offene Fragen konzipiert, aus denen Erkenntnisse über die gewählte Strategie abgeleitet werden sollen. Mit Hilfe qualitativer Inhaltsanalyse (Mayring, 2015) wird ein Kategoriensystem angestrebt, das verschiedene Bewältigungstypen abbildet und Aufschlüsse über innere Prozesse der Teilnehmer*innen gibt. Anschließend wird eine Quantifizierung der Typen vorgenommen und der Zusammenhang zwischen Selbstkonzeptveränderung und Bewältigungsstrategie analysiert. Demzufolge soll gezeigt werden, ob bestimmte Typen höhere Veränderungen im Selbstkonzept aufweisen als andere, um dadurch Rückschlüsse auf die Effektivität der Bewältigungsstrategien ziehen zu können. Es wird erwartet, dass erste Ergebnisse bereits zur Konferenz zur Verfügung stehen.

90. Deskriptive, präskriptive und proskriptive Stereotype über Kinder in Deutschland

Hannah Streck, Ursula Kessels
Freie Universität Berlin, Germany

Geschlechterstereotype enthalten Attribute, die weiblichen und männlichen Personen aufgrund ihres Geschlechtes zugeschrieben werden (Wirtz, 2020). Die Literatur unterscheidet hierbei *deskriptive* Stereotype, die beschreiben, wie weibliche und männliche Individuen sich "typischerweise" verhalten, und *präskriptive* Stereotype, die das für ein Geschlecht „wünschenswerte“ Verhalten beinhalten (Nauts, 2015; Rudman, Moss-Racusin, Phelan & Nauts, 2012; Sullivan, Moss-Racusin, Lopez & Williams, 2018). Andere Ansätze führten als eine Unterform der präskriptiven Stereotype *proskriptive* Geschlechterstereotype ein; diese beschreiben, welche Eigenschaften von weiblichen und männlichen Personen nicht gezeigt werden sollten (Koenig, 2018; Prentice & Carranza, 2002; Sullivan et al., 2018). Geschlechterstereotypen sind für das schulische Engagement von Mädchen

und Jungen höchst relevant (Kessels, Heyder, Latsch & Hannover, 2014), aber für die frühkindliche Bildung liegen noch wenig Befunde vor.

Im Einklang mit Bestrebungen der Gleichstellung der Geschlechter (European Institute for Gender Equality, 2020) berichten die Mehrheit von Erziehungsberechtigten, dass in Kindertagesstätten keine geschlechtsspezifischen Rollenbilder übermittelt werden sollten (Wößmann, Lergetporer, Grewenig, Kersten & Werner, 2018). Jedoch zeigen Studien aus Serbien und den USA, dass Kinder die sich nicht geschlechtstypisch verhalten, von Kindern (Skočajić, Radosavljević, Okičić, Janković & Žeželj, 2020) und Erwachsenen (Sullivan et al., 2018) schlechter angesehen werden als geschlechtstypische Kinder. Für Kinder existieren also durchaus Geschlechterstereotype, die beeinflussen, wie sie handeln und behandelt werden, auch wenn diese sich inhaltlich von Stereotypen über Erwachsene unterscheiden (Koenig, 2018; Sullivan et al., 2018). Stereotype über Kinder beziehen sich laut Literatur eher auf Aktivitäten und Aussehen als auf Eigenschaften (Koenig, 2018) und es lassen sich nur wenige proskriptive Stereotype über Kinder ausarbeiten (Sullivan et al., 2018). Das Ziel unserer Studie ist es, zu prüfen welche Merkmale für Kinder in Deutschland als stereotypisch (deskriptiv, präskriptiv, proskriptiv) gelten.

Methode:

In früheren Studien wurde zur Unterscheidung in deskriptive, präskriptive und proskriptive Stereotype folgendermaßen vorgegangen (Rudman, Moss-Racusin, Phelan, et al., 2012; Sullivan et al., 2018; Nauts, 2015; Koenig, 2008): Versuchspersonen bewerteten auf einer neunstufigen Likertskala, welche Merkmale für Jungen bzw. Mädchen als typisch bzw. als wünschenswert gelten. Als „deskriptiven“ Stereotypen entsprechend wurden jene Merkmale ausgewählt, die bzgl. Typizität bei einem Geschlecht einen Wert > 6 erhielten und zudem einen signifikanten Unterschied mit einer Effektgröße über $|d| = 0,4$ zum anderen Geschlecht aufwiesen. Mit dem gleichen Vorgehen wurden „präskriptive“ Items bzgl. der Erwünschtheit ausgewählt. Darüber hinaus wurden Items, die auf der Skala für wünschenswertes Verhalten mit < 4 bewertet wurden und signifikante Geschlechtsunterschiede mit mind. $|d|= 0,4$ haben, als „proskriptiv“ eingestuft.

Unsere Studie folgte diesem Vorgehen, um zu erfassen, welche deskriptiven, präskriptiven und proskriptiven Geschlechterstereotype für Kinder in Deutschland gelten. Dafür wurden N = 390 Studierende (250 weiblich, 104 männlich, 12 divers/trans*/non-binär, 24 keine Angabe) aus mehreren Fachbereichen rekrutiert. In einem 2x2 between-subjects design (Geschlecht Target: Mädchen, Junge; Rating: typisch, wünschenswert) wurden 72 items (36 weiblich, 36 männlich konnotiert) auf einer 9 Pkt. Likertskala ausgewertet. Teilnehmer/innen wurden gebeten zu bewerten, ob die Merkmale in der deutschen Gesellschaft als typisch (bzw. wünschenswert) für Jungen (bzw. Mädchen) gelten (1 = „gar nicht typisch“, 9 = „sehr typisch“ [bzw. „wünschenswert“]).

Erste Ergebnisse zeigten, dass sich auch in Deutschland deskriptive, präskriptive und proskriptive Geschlechterstereotype für Kinder existieren. Ähnlich wie Sullivan et al. (2018) und Koenig (2018) zeigen wir, dass Items zu Aussehen und Aktivitäten („spielt mit Puppen“) als stereotypischer gelten als Eigenschaften („durchsetzungsfähig“). Zudem sind mehr proskriptive Stereotype über Kinder festzustellen als in einer ähnlichen Studie aus den USA (Sullivan et al., 2018).

Unsere Ergebnisse sollen die Grundlage dafür bilden, im deutschen Raum genauer zu untersuchen, wie Kinder, die gegen deskriptive, präskriptive oder proskriptive Geschlechterstereotypen verstoßen, von anderen Sozialisationspersonen bewertet werden. Zudem wird die theoretische Unterteilung in deskriptive, präskriptive und proskriptive Stereotype anhand der Daten kritisch hinterfragt.

91. Schutzfaktoren von Kindern und Jugendlichen in Zeiten von (Teil-)Schulschließungen, Isolation und Quarantäne

Eva Grommé, Christian Reintjes, Sonja Nonte
Universität Osnabrück, Germany

Die Pandemie ist seit zweieinhalb Jahren ein anhaltendes Thema. Die mit den Lockdowns verhängten *Non-Pharmaceutical Interventions* wie (Teil-) Schulschließungen sowie die Minimierung sozialer Kontakte haben tief in Organisation und Ablauf von Schule und Unterricht (Reintjes, Porsch & Brahm, 2021), aber auch in die Freizeitgestaltung der Kinder und Jugendlichen (Wößmann et al., 2021) eingegriffen. Nach den Schulschließungen hat sich inzwischen eine Phase der individuellen und stellenweisen Hygieneschutzmaßnahmen eingestellt (Reintjes & Brahm, 2021). Der mit Quarantäne- und Isolationsmaßnahmen einhergehende eingeschränkte Zugang zur Schule birgt dabei andere Herausforderungen als eine wochenlange kollektive Abwesenheit während den flächendeckenden Schulschließungen (Budde et al., 2022). Neben schul- und unterrichtsbezogenen Folgen der Pandemie auf kognitiver Ebene belegen Forschungsarbeiten (negative) Konsequenzen für das Befinden, die (mentale) Gesundheit sowie die Lebensqualität von Kindern und Jugendlichen (vgl. hierzu ausführlich Reintjes et al., 2021).

Theoretisch ist davon auszugehen, dass Belastungen von Kindern und Jugendlichen durch Schutzfaktoren abgemildert werden können (Wieland, 2011). Solche Schutzfaktoren sind entweder persönlich veranlagt oder werden durch die Umwelt gestellt, wie beispielsweise familiale soziale Unterstützung (Werner & Smith, 2001). Als umweltbezogene Risikofaktoren gelten u.a. Armut, ein niedriger Bildungsstand der Eltern oder fehlende soziale Netzwerke (Oerter et al., 2011). Zudem sollten Schutzfaktoren situationsbedingt und altersabhängig betrachtet werden (Wyman et al., 1999). Im Kontext der Schulschließungen zu Beginn der Pandemie konnten insbesondere soziale Kontakte als wichtige Ressource herausgestellt werden (Cohen et al., 2022; Engelke et al., 2022; Gayatri & Irawaty, 2022), wohingegen eine geringe Bildung der Eltern, wenig monetäre Ressourcen oder auch eingeschränkter Wohnraum als Risikofaktoren gelten (Anders et al., 2021; Lips et al., 2022).

Aus bildungspolitischer Sicht wird eine zwischen den Bundesländern, aber auch zwischen den Städten und Gemeinden stark variierende Aufholstrategie (Aufholen nach Corona) avisiert (KMK, 2022). Das kommunale Bildungsbüro der Stadt Osnabrück entschied sich im Frühjahr 2021 für ein im Kontext der Aufholstrategie verortetes datengestütztes Monitoring, unter anderem des subjektiven Wohlbefindens der Schüler:innen, der Zufriedenheit mit der Schule, des digitalen Lernens und der Zufriedenheit mit den verfügbaren Freizeitangeboten. Im Juni 2021 und Mai 2022 wurden mit Hilfe einer Online-Befragung ("Muntermacher", Reintjes et al., 2022) Informationen hierzu erhoben. Die Ergebnisse wurden öffentlich präsentiert und diskutiert. Im Mai 2022 beschloss auch Mülheim a.d.R., ein derartiges Monitoring durchzuführen ("Schule und Corona", SchuCo).

Ausgehend davon, dass mit individuellen Fehlzeiten aufgrund von Corona erweiterte Belastungen bei Kindern und Jugendlichen einhergehen, werden persönliche, familiale und schulische Schutzfaktoren für den Umgang mit Quarantäne und Isolation untersucht (Grommé et al., 2023). Hierzu wurden zunächst die Datensätze aus Mülheim (*n*=923) und Osnabrück (*n*=576) zusammengeführt. Zudem wurden die Kinder und Jugendliche von den weiteren Analysen ausgeschlossen, die angaben, sich noch nie in Quarantäne oder Isolation befunden zu haben (*n*=411). Anschließend wurden mit *Mplus 8.6* (Muthén & Muthén, 1998-2017) multiple lineare Regressionsanalysen getrennt nach Grund- und Sekundarstufe durchgeführt, wobei ein robuster Schätzparameter (MLR) eingesetzt und fehlende Werte durch eine Multiple Imputation geschätzt wurden (*m*=10). Die Stichprobenzugehörigkeit (Cluster) Mülheim/Osnabrück wurde anhand des Befehls type=complex kontrolliert.

Die Ergebnisse zeigen, dass Grundschüler:innen besser mit Quarantäne und Isolation umgehen können, wenn sie sich zu Hause sozial unterstützt fühlen, die Menge an Fehltagen geringer ist und sie jünger sind. Bei den Sekundarschüler:innen zeigen sich positive Zusammenhänge, wenn sie sich generell psychisch und körperlich wohl fühlen, aber auch wenn sie ein eigenes Zimmer haben. Zudem gaben Mädchen häufiger an, schlechter mit der Quarantäne umgehen zu können als Jungen.

Resümierend zeigen sich Schnittmengen mit Faktoren, die im Alltag von Kindern und Jugendlichen schon vor der Pandemie wichtig waren und in bildungspolitischen Aufholprogrammen auch im Kontext von Corona eine tragende Rolle spielen sollten. Hierbei dient regionales Bildungsmonitoring als passendes Instrument, um eine evidenzbasierte Diagnose von Bedarfen einzelner Schulen zu ermöglichen und konkrete Unterstützungsmaßnahmen zu implementieren.

92. Rekonfigurationen der Sozialstruktur und das Impostor Phänomen – Ein fundiertes sozialwissenschaftliches Theoriekonzept und ein innovatives empirisches Forschungsdesign

Max Jansen
Goethe-Universität Frankfurt, Deutschland

Positionen innerhalb einer sozialen Struktur verschaffen Einzelpersonen und Gruppen einen unterschiedlichen Zugang zu sozialem Status und Macht.[i] Weil demokratische Gesellschaften in Bezug auf die Zuweisung sozialer Positionen in besonderem Maße leistungsorientiert sind, gilt Bildung hier als zentrale Ressource und eine der vielversprechendsten Triebkräfte für sozialen Aufstieg.[ii] In den letzten Jahrzehnten kam es durch die Bildungsexpansion, berufliche Aufwertungen, steigende Frauenerwerbstätigkeit und Migration zu massiven sozio-strukturelle Rekonfigurationen, die die soziale Mobilität vormals marginalisierter Gruppen erhöht haben.[iii]

Derartige soziostrukturelle Verschiebungen können Individuen mit Situationen konfrontieren, die die Überzeugungen, Normen und selbstverständlichen Handlungen ihrer Herkunftsgruppe infrage stellen, da sie neue Erfahrungen hervorbringen, die Einzelpersonen dazu veranlassen, ihre Positionen und Handlungen neu zu bewerten.[iv] Mit einem sozialen Aufstieg zusammenhängende Statusinkonsistenzen können sozialen Druck erzeugen, weil an Individuen uneinheitliche, manchmal gar widersprüchliche Anforderungen und Erwartungen gestellt werden.[v] Ein Phänomen, das in diesem Zusammenhang aktuell ungekannte Popularität erfährt, ist das Impostor Phänomen (IP, d.h. Hochstapler Phänomen). Jährliche werden über 1.000 akademische Beiträge zum IP publiziert, während die sozialen Medien und die Ratgeber- und Selbsthilfeliteratur in ähnlichem Maße floriert.

Seitdem die Psychologinnen Pauline Rose Clance und Suzanne Imes dem Impostor Phänomen 1978 seinen Namen gaben,[vi] wurde facettenreich beschrieben, wie sich Menschen fühlen, die ihre Bildungs- und Arbeitsmarkterfolge nicht ihren eigenen Fähigkeiten, sondern instabilen äußeren Faktoren zuschreiben: Sie glauben, ihre persönlichen Errungenschaften guten Kontakten, einem guten Timing, ihrem Charme oder gar reinem Glück oder dem Zufall zu verdanken, haben das Gefühl, von anderen überschätzt zu werden, Zweifeln an ihren eigenen Fähigkeiten und haben Angst, als „Hochstapler:in“ entlarvt zu werden. Heute fällt auf, dass die Forschung es bislang versäumt hat, den sozialen Wurzeln dieses Phänomens ausreichend theoretische und empirische Aufmerksamkeit zu widmen, obwohl Clance et al. (1995) betonten, dass IP Gefühle maßgeblich von „interpersonellen und sozialen Kontexten“ geprägt sind.[vii]

Stattdessen beschränkt sich die Mehrzahl bestehender Forschungsartikel auf die Untersuchung individueller Faktoren und spart Kontextvariablen auf gesellschaftlicher, institutioneller und zwischenmenschlicher Ebene systematisch aus.[viii] Hierbei spielen die vorherrschende Konzeptualisierung des IP als individuelle Pathologie sowie die gängigen empirischen Forschungspraxen eine zentrale Rolle. Vorhandene Befunde basieren fast ausschließlich auf Querschnittdesigns mit sehr kleinen und hochselektiven Stichproben und sind durch große Defizite hinsichtlich der Dokumentation wesentlicher demografischer Variablen gekennzeichnet.[ix]

Vor dem Hintergrund, dass IP Gefühle bereits in der späten Grundschulzeit zu beobachten sind,[x] das junge Erwachsenenalter einen Zeitraum darstellt, in dem das Risiko, IP zu entwickeln, seinen Höhepunkt erreicht[xi] und Diskrepanzen zwischen selbstwahrgenommenen und fremd zugeschriebenen Kompetenzen in diesem Alter als unwiderlegbarer Beweis für die eigene Inauthentizität interpretieren werden,[xii] stellt dieses Poster eine bildungs- und erziehungssoziologisch fundierte Theorie zur Entstehung und Genese des IP sowie ein innovatives Messinstrument zu dessen empirischer Längsschnittuntersuchung vor.

Aufbauend auf die *social identity theory* [xiii] und die *ecological systems theory* [xiv] nimmt der hier vorgestellte theoretische Rahmen Einflüsse von Kontextfaktoren ab dem Kindheits- und Jugendalter in den Blick und berücksichtigt insbesondere die Tatsache, dass *Statusinkonsistenzen* in den letzten Jahrzehnten massiv zugenommen haben.[xv] Damit verbunden wird die Entwicklung, Validierung und Implementierung der *Impostor Phenomenon Short Scale* (IPSS) vorgestellt, die es erlaubt, die Genese und Entwicklung des IP ab dem frühen Jugendalter empirisch zu erfassen. Mithilfe der IPSS wird ab März 2023 die weltweit erste registerbasierte Stichprobenerhebung zum IP durchgeführt. Hierzu wird ein repräsentatives Sample von 15-17-Jährigen (n = 4000) erhoben. Die detaillierte Darstellung des theoriebasierten Entwicklungsprozesses der IPSS sowie die Ergebnisse der Validierungsstudien dienen als best practise Beispiele und tragen dazu bei, den sozialwissenschaftlichen Forschungsstand zum IP maßgeblich zu erweitern und dabei die Standards guter wissenschaftlicher Praxis in Bezug auf das IP und die umfassende Dokumentation psychometrischen Daten zu forcieren.

93. Berufliche Orientierung in unsicheren Zeiten - Erste Befunde aus dem Modellvorhaben "VerOnika"

Ariane Neu
FernUniversität in Hagen, Deutschland

Das Modellvorhaben „VerOnika – Verzahnte Orientierungsangebote zur beruflichen und akademischen Ausbildung“ zielt auf die Entwicklung und Erprobung hybrider Orientierungsangebote. Diese Orientierungsangebote wenden sich an hochschulzugangsberechtigte Personen und sollen den Teilnehmenden gleichwertig Einblicke in ein Hochschulstudium sowie in eine berufliche Ausbildung bieten, um so eine erfahrungsbasierte Bildungsentscheidung zu ermöglichen. Entwickelt und erprobt werden diese verzahnten Orientierungsangebote an den drei Standorten Berlin, Darmstadt und Karlsruhe jeweils in Zusammenarbeit von Hochschulen der angewandten Wissenschaft und Partnern der beruflichen Bildung. Das Lehrgebiet Lebenslanges Lernen der FernUniversität in Hagen hat im Rahmen dieses BMBF-geförderten Verbundvorhabens die wissenschaftliche Begleitung übernommen (Projektlaufzeit: 12/2019 bis 09/2023). Ziel der wissenschaftlichen Begleitung ist es, prozessbezogen die erfolgreiche Entwicklung und Erprobung der verzahnten Orientierungsangebote durch empirische Untersuchungen zu unterstützen und abschließend weitergehende Handlungsempfehlungen für die Entwicklung und Durchführung verzahnter Orientierungsangebote an der Schnittstelle zwischen beruflicher und akademischer Bildung abzuleiten. Als zentrale empirische Erhebungsmethoden kommen dazu u. a. Dokumentenanalysen, leitfadengestützte offene Expert*inneninterviews, problemzentrierte Gruppeninterviews mit Teilnehmenden der Orientierungsangebote sowie quantitative Online-Befragungen der Teilnehmenden zum Einsatz.

Der Beitrag stellt erste Befunde dieser wissenschaftlichen Begleitung vor, die sich insbesondere auf die Teilnehmenden-Perspektive beziehen. So werden Ergebnisse aus den problemzentrierten Gruppeninterviews präsentiert, die in Anlehnung an Witzel (1982, 1985) jeweils zu Beginn der Orientierungsprogramme mit den Teilnehmenden durchgeführt werden. Thematisiert werden in diesen Gruppeninterviews u.a. die Fragen, warum sich die Teilnehmenden für das Orientierungsprogramm entschieden haben und welche Ziele sie mit ihrer Teilnahme verfolgen. Des Weiteren wird besprochen, welche Bedürfnisse und Erwartungen die Teilnehmenden an die Orientierungsprogramme haben und wo sie besondere Herausforderungen in der Gestaltung ihres nachschulischen Bildungsweges sehen. An diesen Gruppeninterviews haben bislang 83 Personen teilgenommen (Stand: 14. Oktober 2022). Die Interviews wurden aufgrund von Corona digital via Videokonferenzsystem durchgeführt, aufgezeichnet und anschließend vollständig transkribiert. Die Auswertung der Interviews erfolgt mittels qualitativer Inhaltsanalyse nach Mayring (2015), wobei eine Kombination von deduktiven und induktiven Verfahrensweisen zum Einsatz kommt.

Ergänzend zu diesen Gruppeninterviews präsentiert der Beitrag erste Befunde aus den quantitativen Online-Befragungen der Teilnehmenden, die jeweils zum Abschluss der Orientierungsprogramme durchgeführt werden und an denen bislang 36 Personen teilgenommen haben (Stand: 14. Oktober 2022). In diesen Abschlussbefragungen werden u.a. die Erfahrungen angesprochen, die die Teilnehmenden mit dem jeweils besuchten Orientierungsprogramm gesammelt haben sowie die Frage, welche Elemente des Orientierungsprogrammes sie für besonders hilfreich für ihren individuellen Orientierungsprozess empfunden haben. Des Weiteren werden sie um Einschätzungen zu verschiedenen Bildungswegen gebeten und ihre Pläne für ihren weiteren Bildungsweg abgefragt.

Um u.a. der Frage nachgehen zu können, ob bestimmte Personengruppen besonders von den verzahnten Orientierungsprogrammen angesprochen werden und bestimmte Programmelemente als besonders hilfreich empfinden, werden darüber hinaus auch zentrale soziodemografische Daten erhoben.

Bei einer Bewertung dieser Befunde ist zu berücksichtigen, dass die konkrete Umsetzung der Orientierungsangebote fast zeitgleich mit der Corona-Pandemie begann. Entsprechend standen nicht nur die institutionellen Akteure dieser Programme vor neuen, bisher unbekannten Herausforderungen, sondern auch die Teilnehmenden waren und sind durch die Corona-Pandemie sowie seit diesem Jahr ergänzend durch die Unsicherheiten des Ukraine-Krieges mit bislang unbekannten wirtschaftlichen und gesamtgesellschaftlichen Rahmenbedingungen konfrontiert. So konnten bspw. in Folge der Corona-Pandemie Praxisphasen in Unternehmen nicht immer im ursprünglich geplanten Umfang durchgeführt werden.