

Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens

- Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I -

Markus Emden

Gefördert durch: SCHERING STIFTUNG Deutsche Forschungsgemeinschaft DFG

Theoretischer Hintergrund

Der Prozess naturwissenschaftlich-experimentellen Arbeitens:

- ⇒ Experimentieren: Kompetenz im Bereich ‚naturwissenschaftliche Erkenntnisgewinnung‘ (KMK, 2005a-c)
- ⇒ Kompetenzen, ergo Schülerleistung im Experimentieren, sollen prinzipiell empirisch überprüfbar sein (Klieme et al., 2003)
- ⇒ Vorgeschlagene Sequenzierungen (s. u.) des Experimentierens meist parallel zu einem vereinfachten, SDDS-analogen Dreischritt (=NAW): [1] Idee/Hypothese formulieren, [2] Experiment planen/durchführen, [3] Schlussfolgern
- ⇒ MC-Test zur Prüfung des Dreischritts: NAW-Test (Klos et al., 2008)

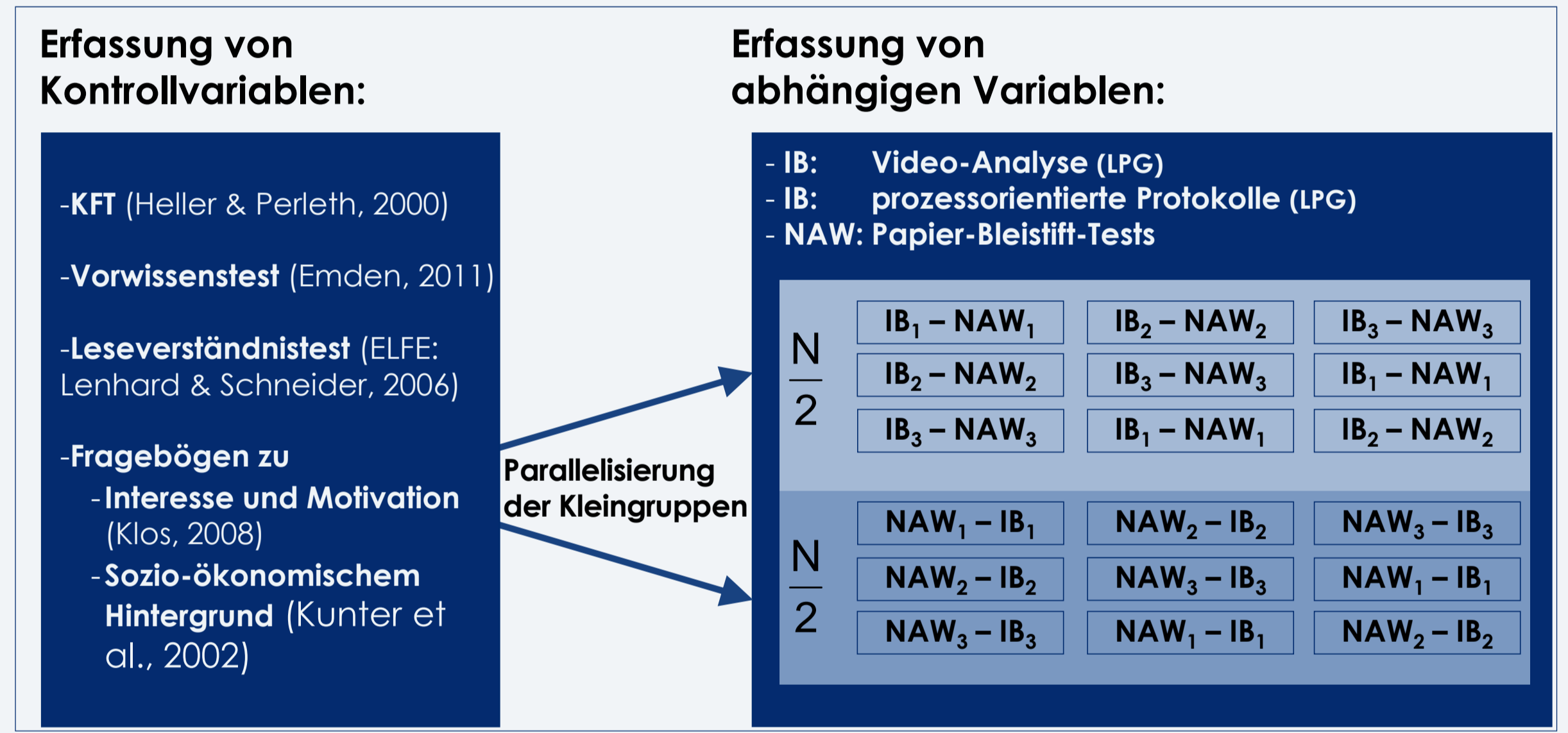
	[1]	[2]	[3]
Schreiber, Theyßen, & Schecker (2009)	Planung	Durchführung	Auswertung
Harms & NAW (2006)	Frage/Problem/Hypothese aufwerfen	Untersuchung planen und durchführen	Be- und Auswertung Hypothesenrevidieren
Kipnis & Hofstein (2009)	Problem-identifikation Hypothesenformulierung	Planung Daten sammeln und auswerten	Schlussfolgerung
Mayer (2007)	rw. Fragen formulieren Hypothesen generieren	Untersuchungen planen	Daten analysieren/ Schlussfolgerung
Hammann (2004), Walpuski (NAW, 2006)	Idee/ Hypothese finden	Durchführung	Schlussfolgerung
Zimmerman (2000)	Hypothesengenerierung	Experimentieren	Evidenz-evaluation Hypothesenrevidieren
Lunetta (1998)	Planung und Design	Durchführung	Analyse und Interpretation
Davies (1996)	Findung eines Lösungsweges	Planung und Durchführung	Auswertung
Khalil & Dunbar (SDDS, 1988)	Suchen im Hypothesenraum	Prüfen von Hypothesen	Evaluieren von Evidenz
Kempa (1986)	Problemformulierung	Entwurf/ Planung Aufbau/ Durchführung Messen/ Beobachten	Interpretation

Performance Assessments:

- ⇒ Empirische Prüfung mittels Realexperimenten beeinträchtigt durch zeitlichen, organisatorischen und personellen Aufwand (z. B. Garden, 1999)
- ⇒ Vorschläge alternativer Erhebungsmethoden favorisieren offene Formate (z. B. Baxter & Shavelson, 1994)
- ⇒ Möglichst: Erhebung durch mehrere Aufgaben (z. B. Gott & Duggan, 2002)
- ⇒ Möglichst: Berücksichtigung mehrerer Messzeitpunkte (Cronbach et al., 1997; Shavelson et al., 1999)
- ⇒ Bisher keine empirische Prüfung alternativer Maße basierend auf einheitlich definierter Prozessstruktur wie z. B. SDDS oder NAW

Forschungsfragen und Design

- FF 1: Gelingt es mit einer offenen, an SDDS/NAW orientierten Protokollmethode Schülerleistung in Experimentierprozessen valide zu erheben?
- FF 2: Gelingt es mit einem Multiple-Choice-Test (NAW-Test) Schülerleistung in Experimentierprozessen valide zu erheben?
- H1/H2: Korrelationen $\geq .65$ zwischen den Leistungsmaßen und einem Maß aus Videoanalysen sprechen für Validität
- Instrumente zur Erhebung der Schülerleistung in Experimentierprozessen:
 - ⇒ Drei offene Experimente in Kleingruppen (IB: Sumfleth, Rumann & Nicolai, 2004)
 - ⇒ Begleitend: an SDDS/NAW prozessorientierte Protokolle (Kemper & Tatzki, 2010)
 - ⇒ Drei thematisch angebundene NAW-Tests (Emden, 2011)
 - ⇒ Video- und Protokollanalyse durch SDDS-analoge Kodierung in Lernprozessgrafiken (LPG: Walpuski & Sumfleth, 2009)
 - ⇒ Generierung eines Leistungsmaßes für den experimentellen Prozess aus LPG (Kemper & Tatzki, 2010)



Zentrale Ergebnisse

Prüfung des Prozesses naturwissenschaftlich-experimentellen Arbeitens mit alternativen Maßen:

⇒ Rangkorrelationen Video-/Protokollmaß über alle Messzeitpunkte (MZP)

	r_s	Gesamt			Gymnasium			Gesamtschule		
		IB 1	IB 2	IB 3	IB 1	IB 2	IB 3	IB 1	IB 2	IB 3
Video/Protokoll	Sig. (2-seitig)	.492***	.521***	.383**	.801***	.861***	.294	.348	.300	.510**
	n	.000	.000	.001	.000	.000	.097	.060	.076	.001
		60	60	69	30	24	33	30	36	36

⇒ Rangkorrelationen Video-/Protokollmaß nach MZP differenziert

		MZP 1	MZP 2	MZP 3
		Gymnasium	IB 1	.442
	Video/Protokoll	.234	.000	.053
		9	10	6
	IB 2	1.000**	1.000**	.789**
	Video/Protokoll	.4	.5	.12
		4	5	12
Gesamtschule	IB 1	.224	.028	.825**
	Video/Protokoll	.562	.943	.006
		9	9	9
	IB 2	-.357	.225	.866*
	Video/Protokoll	.192	.532	.026
		15	10	6

MZP 1 → MZP 3
Stabilisierung von r_s bei hoher Stärke

⇒ **Gewöhnung an Messmethoden notwendig**
⇒ **Extraneous Load (Sweller, 2010) bei Koordinierung der Methoden beeinträchtigt die beobachtbaren Zusammenhänge**

MZP 1 → MZP 3
Gegenläufige Entwicklung von r_s für die beiden Schulformen

Gy → GeS
⇒ **wachsende Gewöhnung führt für Gy zu Expertise Reversal Effect (Kalyuga et al., 2003)**
⇒ **Schwierigkeit der Aufgabe unterschiedlich angemessen für Leistungs-niveaus**

⇒ Rangkorrelationen Video-/NAW-Maß, über alle MZP

	r_s	Gesamt	Gymnasium	Gesamtschule
		IB 1 Video/NAW	.313*	.491**
	Sig. (2-seitig)	.015	.006	.283
	n	60	30	30

Ausschluss NAW zu IB 2 und IB 3 aufgrund unzureichender Testgüte (Cronbachs Alpha)
 r_s erlaubt keine Bestätigung der Hypothese zu FF 2

Folgerungen und Perspektiven

Folgerungen:

- Aus FF1:** Prozessorientiertes Protokollieren zeigt prinzipiell Eignung zur Bewertung von Experimentierprozessen ($r_s > .65$), vorausgesetzt:
 - ⇒ Lernende haben Möglichkeit sich an das Format zu gewöhnen,
 - ⇒ Themenstellung des Experiments ist auf die Leistungsfähigkeit abgestimmt.
- Aus FF2:** Das geschlossene Format zeigt aufgrund geringer Testgüte keine gute Eignung zur Bewertung von Experimentierprozessen ($r_s < .65$).

Perspektiven:

- ⇒ Anschlussstudien sollten den *Cognitive Load* bei der Bearbeitung von Experimenten und Protokollen berücksichtigen, um:
 - ⇒ Gewöhnungseffekte zu bestätigen,
 - ⇒ Einfluss unterschiedlich schwieriger Aufgabeninhalte weiter zu untersuchen.
- ⇒ Weitere Aufklärung der Eignung von NAW-Tests sollte durch bewährte Tests mit höherer Testgüte erfolgen.

Literatur

Baxter, G. P., & Shavelson, R. J. (1994). Science Performance Assessment, Benchmarks and Sumogates. *International Journal of Educational Research*, 27 (2), 279-298.

Cronbach, L. J., Linn, M. C., Brennan, R. L., & Haertel, E. H. (1997). Generalizability Analysis for Performance Assessments of Student Achievement or School Effectiveness. *Educational and Psychological Measurement*, 57 (3), 373-399.

Dowse, E. (1998). In Search of Sources: Some Issues for Chemists and Other Problem-Solving Activities in Chemistry. *London: Royal Society of Chemistry*.

Emden, M. (2011). *Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens: Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*. Berlin: Logos.

Garden, R. A. (1999). Development of TMS Performance Assessment Tasks. *Studies in Educational Evaluation*, 25 (2), 217-241.

Gott, R., & Duggan, S. (2002). Problems with the Assessment of Performance in Practical Science Which Way Now? *Cambridge Journal of Education*, 32 (2), 183-201.

Hammann, M. (2004). *Kompetenzentwicklungsmessung: Merkmale und ihre Bedeutung: dargestellt anhand von Kompetenzen beim Experimentieren. Der mathematische und naturwissenschaftliche Unterricht (MNU)*, 57 (4), 196-203.

Heller, K. A., & Perleth, C. (2003). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen (KFT 4-12/0)*. Göttingen: Hogrefe Verlag für Psychologie.

Kempes, S., Ajzen, F., Chander, P., & Sweller, J. (2010). The Expertise Reversal Effect. *Educational Psychology*, 38 (1), 23-31.

Kempa, R. (1986). *Assessment in Science: Cambridge Science Education Series*. Cambridge: Cambridge University Press.

Kemper, S., & Tatzki, V. (2010). Entwicklung eines Bewertungsschemas zur Leistungsbeurteilung von Lernprozessgraphiken und Protokollen. *Universität Duisburg-Essen: Unveröffentlichte Seminararbeit im Modul 'Wissenschaftliches Arbeiten'*.

Kipnis, M., & Hofstein, A. (2009). The Inquiry Laboratory as a Source for Development of Metacognitive Skills. *International Journal of Science and Mathematics Education*, 37 (1), 401-427.

Klos, S., & Zentgraf, E. (1998). Dual Source Search During Scientific Reasoning. *Cognitive Science*, 22 (1), 148.

Klieme, E., Avenarius, H., Blum, W., Dörmich, P., Gruber, H., Prenzel, M., et al. (2003). *Jur. Entwicklung nationaler Bildungsstandards: Expertise: Bildungsforschung*. Bonn: Bertelsmann.

Klos, S. (2008). *Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht: Der Einfluss eines integrierten Unterrichtsansatzes*. Berlin: Logos.

Klos, S., Henke, C., Klieme, E., Walpuski, M., & Sumfleth, E. (2008). Naturwissenschaftliches Experimentieren und chemisches Fachwissen - zwei verschiedene Kompetenzen. *Zeitschrift für Pädagogik*, 54 (3), 304-321.

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., et al. (2002). *FSA 2002: Dokumentation der Erhebungsinstrumente: Materialien aus der Bildungsforschung*. Berlin: Max Planck Institut für Bildungsforschung.

Lenhard, W., & Schneider, W. (2006). *ELFE 1-4: Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe Verlag für Psychologie.

Lunetta, V. A. (1998). The School Science Laboratory: Historical Perspectives and Context for Contemporary Teaching. In H. J. Rose & G. Tobin (Eds.), *International Handbook of Science Education* (S. 249-262). Dordrecht u. a.: Kluwer.

Mayer, J. (2007). *Erkenntnisgewinnung aus wissenschaftlichen Problemstellungen*. In D. Krüger & H. Vogt (Eds.), *Prozesse in der kognitiven Psychologie* (S. 177-186). Berlin, Heidelberg, New York: Springer.

Schreiber, H., Theyßen, H., & Schecker, H. (2009). Experimentelle Kompetenz messen? *Physik und Didaktik in Schule und Hochschule*, 8 (3), 92-101.

Selektoral der 83rdigen Konferenz der Kulturlenker der Länder in der Bundesrepublik Deutschland (KMK). (2005). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.

(2005). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.

(2005). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. Beschluss vom 16.12.2004. München: Luchterhand.

Shavelson, R. J., Ruiz-Primo, M., Aronak, & Wiley, E. W. (1999). Note on Sources of Sampling Variability in Science Performance Assessment. *Journal of Educational Measurement*, 36 (1), 61-71.

Sumfleth, E., Rumann, S., & Nicolai, R. (2004). *Kognitives Arbeiten in Chemieunterricht*. Essener Unikate, 24, 75-85.

Sweller, J. (2010). *Benevolent Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load*. *Educational Psychology Review*, 22 (2), 123-138.

Walpuski, M. (2006). *Optimierung von experimenteller Arbeitsgruppenarbeit durch Strukturgraphiken und Feedback*. Berlin: Logos.

Walpuski, M., & Sumfleth, E. (2009). The Use of Video Data to Evaluate Inquiry Studies in Chemistry Education. In J. Lank & S. Edel (Eds.), *The Power of Video Studies in Investigating Teaching and Learning in The Classroom* (S. 121-133). Münster: Waxmann.

Wissenschaftliches Konstrukt im Harms'schen Naturwissenschaftstest - Kompetenzmodell und Konzepte für Bildungsstandards. *Berlin (Wissenschaftlicher Schulabschluss)*.

Zimmerman, C. (2000). *The Development of Scientific Reasoning Skills*. *Developmental Review*, 20 (3), 99-149.

Kontakt



Markus Emden
Universität Duisburg-Essen
Forscherguppe und Graduiertenkolleg
'Naturwissenschaftlicher Unterricht'
Schützenbahn 70, 45127 Essen
Markus.Emden@uni-due.de