

Von FIMS und FISS bis TIMSS und PISA

Schulleistungen in Deutschland im historischen und internationalen Vergleich

Zu Beginn des 21. Jahrhunderts wird Deutschland mit einem bitteren Befund konfrontiert: *PISA*, die groß angelegte internationale Vergleichsstudie, bescheinigt den Schulen des Landes, dass sie ihre SchülerInnen am Ende der Schulpflichtjahre auf ihr Leben als Erwerbstätige wie als BürgerInnen schlecht vorbereitet. Angesichts dieser deprimierenden Diagnose ist ein *Blick zurück* reizvoll – nicht zuletzt deshalb, weil er auf ein Kontrastbild fällt: von der internationalen *Spitzenstellung* des deutschen Bildungssystems in den ersten Jahrzehnten bis hin zur *Mittelmäßigkeit* am Ende des 20. und zu Beginn des 21. Jahrhunderts. Die Darstellung wesentlicher Hintergründe und Weichenstellungen im Kontext der Leistungsfähigkeit des deutschen Schulsystems wird in diesem Beitrag durch die Betrachtung der Leistungsergebnisse deutscher SchülerInnen bei internationalen Vergleichsstudien ergänzt. Neben den aktuellen Tests stehen auch die Large Scale Assessments der 60er und 70er Jahre, an denen Deutschland beteiligt war, im Fokus. Sie spielen in der aktuellen Diskussion so gut wie keine Rolle, obwohl der weite Blick zurück – zum Teil auf einer empirischen Basis – *Entwicklungstrends* verdeutlicht. Die Rückschau auf die nationalen Entwicklungslinien wird durch eine *internationale Perspektive* ergänzt: Ein Blick auf die deutschen Ergebnisse und die Resultate wichtiger Vergleichsländer bei internationalen Studien – von FIMS und FISS bis TIMSS und PISA – stellt das deutsche Abschneiden bei Leistungsvergleichen wie PISA in einen weiteren Kontext und liefert Ansatzpunkte für weitere Forschungsaktivitäten und Handlungskonsequenzen.

1. Entwicklungslinien am Beginn des 20. Jahrhunderts

Zu Beginn des 20. Jahrhunderts galt das deutsche Bildungssystem den nationalen wie den internationalen Beobachtern als vorbildlich. Deutschland verfügte damals über *eines der leistungsfähigsten Bildungssysteme*. Das war das Ergebnis von Weichenstellungen, die z.T. schon vor und z.T. nach dem ersten Weltkrieg vorgenommen worden waren:

- Die *Schulpflicht* – im 18. Jahrhundert immer wieder verkündet – war am Ende des 19. Jahrhunderts auch tatsächlich durchgesetzt, ihre Verankerung auch für die Jahre der beruflichen Bildung in der Weimarer Verfassung war wegweisend.
- Die *gymnasiale Bildung war curricular modernisiert*: Der ‚Allerhöchste Erlass‘ von 1900 hatte neben das bis dahin allein abiturführende Altsprachliche Gymnasien die naturwissenschaftlich bzw. neusprachlich geprägten Jungengymnasien gestellt. Mädchen konnten in Deutschland, in Preußen seit 1908, in den Lyzeen ein Abitur erlangen und damit ein Hochschulstudium aufnehmen.

Lehrpläne und Universitäten waren mit ihrer grundsätzlichen Verbindung von Forschern und Lehren, die Humboldt ihnen in die Wiege gelegt hatte, *Innovationsmotoren*. Sie waren dies auch, weil es ein angemessenes Verhältnis zwischen der Zahl der Studierenden und der Lehrenden sowie der Ausstattung der Hochschulen gab.

• Artikel 146 der Weimarer Verfassung und das 1920 folgende „Grundschulgesetz“ schufen mit der *Grundschule als Volksschulunterstufe* die vierjährige gemeinsame Erziehung aller Kinder und damit den Einstieg in eine Schulstruktur, für die nicht länger das Standes-, sondern das *Leistungsprinzip* gültig war.

Dies alles war durch ein international *beachtliches Niveau der Bildungsausgaben* unterfüttert: 1880 brachte Deutschland 1,6 % seines Bruttosozialproduktes für Bildung auf. Der Vergleichswert lag in den USA bei 1,1 %, in Großbritannien und Frankreich bei je 0,9 %. Noch 1930 war Deutschland – trotz des verlorenen Krieges und der ökonomischen Krise im Land – Spitzenreiter: Das Land brachte 4,1 % seines Sozialproduktes für Bildung auf; die Vergleichswerte in den USA und in Großbritannien lagen bei nur 2,8 % (vgl. Maier 1994, S. 46).

Wenn diesem Bildungssystem zwar seine hohe Leistungsfähigkeit attestiert werden kann, nicht aber vorbehaltlos Modernität, so wegen seiner *fehlenden demokratischen Durchdringung*: Der „Weimarer Schulkompromiss“ beschränkte die gemeinsame Erziehung auf die ersten vier Jahre und überließ die Schulkarrieren danach weiterhin der Steuerung mehr durch in der sozialen Herkunft wurzelnde Privilegien denn durch individuell erbrachte Leistungen. Die LehrerInnen der Schulen behielten ihr nach dem Prestige der Bildungsgänge bemessenes gesellschaftliches Ansehen und Selbstbewusstsein, demokratische Überzeugungen prägten durchaus nicht durchgängig die Inhalte des Unterrichts und den Erziehungsstil der Schulen.

Das *frühe Ende der Weimarer Republik* ließ keine Zeit und keinen Raum, dieses Schulsystem auf dem mit der Schaffung der gemeinsamen Grundschule eingeschlagenen Weg weitergehend inhaltlich und strukturell zu modernisieren. An die Stelle einer von demokratischem Denken geprägten Schulpolitik trat die *Schulpolitik der Nationalsozialisten*: Sie füllten die überkommene Schulstruktur mit ihrer demokratiefeindlichen, menschenverachtenden und selektiven Orientierung. Die amerikanische Zook-Kommission, die nach Kriegsende das deutsche Schulsystem begutachtete, konnte daher formulieren: „Dieses System hat bei einer kleinen Gruppe eine überlegene Haltung und bei der Mehrzahl der Deutschen ein Minderwertigkeitsgefühl entwickelt, das jene Unterwürfigkeit und jenen Mangel an Selbstbestimmung möglich machte, auf denen das autoritäre Führerprinzip gedieh“ (zit. nach Herrlitz/Hopf/Titze 1993, S. 161).

2. Entwicklungen seit der Mitte des letzten Jahrhunderts

Anstatt nun nach 1945 den mit dem Ende der Weimarer Republik gerissenen Faden der Modernisierung und Demokratisierung des Bildungssystems wieder aufzunehmen und weiterzuspinnen, ging die *Bundesrepublik* Deutschland den Weg der *Restauration*. Die strukturelle und inhaltliche Wiederherstellung des aus der Weimarer Zeit überkommenen Systems wurde zum Ziel, darüber hinausgehende Ansätze wurden gar nicht erst gewagt oder – wie etwa im west-

lichen Teil Berlins – schnell wieder aufgegeben bzw. im Fall der sechsjährigen Grundschule auf kleinere Veränderungen reduziert. Die *Entwicklung in der DDR*, die mit der Schaffung der Einheitsschule den strukturellen Weg Weimars fortsetzte, diente den Gegnern einer grundsätzlichen Reform der westdeutschen Schulstruktur wegen der engen Verkopplung mit der kommunistischen Ideologie als Abschreckung.

Es waren von außen kommende Beobachter, die 1971 das westdeutsche Bildungssystem im Rahmen eines *OECD-Länderexamens* – schon damals war die OECD Auftraggeber! – begutachteten und ihm ein vernichtendes Urteil ausstellten. Unter dem Titel: „*Bildungswesen: Mangelhaft*“ lautete der Befund: „Die wirtschaftlichen, gesellschaftlichen und politischen Gegebenheiten in der BRD haben sich seit den zwanziger Jahren grundlegend gewandelt. Das Bildungswesen wurde jedoch nach der Hitlerzeit so wieder aufgebaut, wie es vorher gewesen war, und ist in den meisten wichtigen Merkmalen bis heute so geblieben“ (zit. nach Luehrig 1973, S. 46).

Nicht nur die Bedingungen schulischen Lernens wurden damals unter die Lupe genommen. Auch der *messbare Ertrag von Schule*, die Frage, welche Bildungsziele auf welchem Niveau erreicht werden, rückte ins Blickfeld. Zu Beginn der 70er Jahre, als die o.g. OECD-Studie erschien, hatte Deutschland bereits an einer internationalen Leistungsvergleichsstudie teilgenommen. Seit Anfang der 60er Jahre, als erstmals *internationale Large Scale Assessments* durchgeführt wurden, an denen Deutschland schon damals beteiligt war, liegen Daten zu den Leistungen deutscher SchülerInnen vor. Dies ermöglicht ein Nachzeichnen der Leistungsentwicklung des deutschen Schulsystems auf einer empirischen Basis.

2.1 Large Scale Assessments mit deutscher Beteiligung

Bis zu den 1950er Jahren bezog sich die vergleichende Erziehungswissenschaft vor allem auf die *separate Beschreibung der Bildungssysteme in den einzelnen Ländern*. Die schnelle technologische Entwicklung Mitte des letzten Jahrhunderts führte aber schnell zu einem wachsenden Bewusstsein für den *Bedarf einer vorausschauenden Bildungsplanung*, um den rasanten wirtschaftlichen Entwicklungen im internationalen Vergleich gerecht zu werden. In der *vergleichenden Erziehungswissenschaft* wurde die Chance gesehen, die Erfahrungen und Leistungen anderer Länder im Bildungssektor quantitativ aufzuzeigen und sie innovativ für die Evaluation der eigenen Bildungspolitik nutzbar zu machen. International vollzog sich ein Paradigmenwechsel von der Betrachtung schulischen Inputs hin zur *Erfassung des ‚outcome‘ von Schule* – eine Entwicklung, die sich in vielen Ländern auf nationaler Ebene etwa in Form landesweiter Assessments und Examinations fortsetzte. In Deutschland spricht man erst seit Mitte der 1990er Jahre von einem Paradigmenwechsel. Entsprechende *quantitative Methoden*, die einen cross-nationalen Vergleich ermöglichen, wurden entwickelt. Die „*International Association for the Evaluation of Educational Research*“ (IEA) setzte den Startpunkt für die internationalen Vergleiche und entwickelt seit nunmehr über 40 Jahren Large Scale Assessments, in denen schulische Leistungen im Rahmen von Kontext- und Prozessvariablen als Umfeldbedingungen erfasst werden. Dahinter steht bis heute das *Ziel*, ...

1 abelle 1: Organisation, Anlage und Umfang von Large Scale Assessments – Internationale Studien mit deutscher Beteiligung; Stand 2/02 (Isabell van Ackeren; Klaus Klemm)

Kürzel	Koord.	Wiss. Begleit.	Planung	Zeitraum	Ergebnisse	Fachliche Inhalte	Hintergrundmerkmale	Alter/ Stufe	Schulform	N intern. national	N Länder
CivEd	IEA	HU	1993	1995-97; 1999	2001	Politisches Wissen	Pol. Einstellungen, pol. Engagement	14-Jährige	alle ohne Sonders.	90.000	28
Comp Ed	IEA	IPN	1985	1989/90; Phase II: 1992	1993	Computerwissen u. -umgang	Lehrqualifikation, schul. Ausstattung, Lehrpläne, Lehrer- und -fortbildung	5., 8., 11. Klassen	alle ohne Sonderschulen	15-5.000 1152.000	5
FIMS	IEA	DIPF	1961	1963-64	1967	Mathematik	Unterrichtspraxis, Lernumgebung, häusl u. schul. Einfluss auf Leistung	13. Erde S II	Vollk., Real- u. Sonderschulen, Gymnasien	133.000	12
FISS	IEA	DIPF	1966	1970	1973	Nu-Wi; Erdk., Bio., Phy., Methoden	Verhaltensmerkmale, fachliches Interesse, Einstellungen, Unterrichtspraxis	10; 14; Erde S II	Vollk., Real- u. Sonderschulen, Gymnasien	137.000	17 bzw. 19 für ältere Stufen
IRLS	IEA	HU	1988	1990-91	1992; (neu)	Leseverständnis, Lesegewohnheiten	Lesegewohnheiten, familiäre Merkmale, Einstellungen zum Lesen, Unterrichtspraxis	3. und 8. Klassen	alle ohne Sonderschulen	210.000	32
PIRLS/IGLU	IEA	HH	1999	2001	2002/03	Leseverständnis	Leseeigenschaft, Lesemotivation, Freizeitsaktivitäten, professionelle Standards, Unterrichtsprozesse	9- bis 10-Jährige	Grundschulen	210.000	38
PISA	OECD, ACER	MPIB	1998	2000, 2003, 2006	Ende 2001	Leseverständnis, Mathematik, Naturwiss.	fächerübergreifende Kompetenzen, vielfältige Indikatoren für Lernergebnisse und ihre Bedingungen	15-Jährige	alle	5.000 180.000 + 50.000 (PISA-E)	32 + 13 in 2002
TIMSS	IEA	MPIB, IPN, DIPF	1989	1994/95, S II: 1995/96	1996-1998	Mathematik, Naturwiss.	Unterricht, Lehrer, Schulen, außerschulische Lebensumwelt, psychosoziale / individuelle Merkmale	7/8. Kl., Ende S II	alle ohne Sonderschulen	500.000	21 (SI); 16 (SII)
(WCS)	IEA	HH	1981	1984-85	1988	Aufbauleistungen	Aufbaucurriculum, Lehrerverhalten, schulbedingte u. persönliche Faktoren	11. Klassen	allgemein/berufsbildende Schulen	20.000 ca.	14

Abkürzungen:

Kürzel	Name / Bemerkung
ACER	Australian Council for Educational Research
CivEd	Civic Education Study/Citizenship Education
CompEd	Computers in Education Study
DIPF	Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a.M.
FIMS	First International Mathematics Study ;beteiligt: Hessen und Schleswig-Holstein
FISS	First International Science Study
HH	Universität Hamburg
HU	Humboldt-Universität Berlin
IRLS	International Study of Reading Literacy; mit deutscher Erweiterung zu deutschsprachigen Minderheiten in Europa
IEA	Internat. Association for the Evaluation of Educational Achievement
IPN	Institut für die Pädagogik der Naturwissenschaften, Kiel
MPIB	Max-Planck-Institut für Bildungsforschung, Berlin
PISA	Programme for International Student Assessment
PIRLS/IGLU	Progress in International Reading Literacy Study; deutsches Äquivalent: Internat. Grundschul-Lese-Untersuchung
TIMSS	Third International Mathematics and Science Study; Grundschulstudie ohne Deutschland
WCS#?	Written Composition Study; nur Hamburg beteiligt: Hamburger Aufsatzstudie

... Bildungspolitikern und Unterrichtspraktikern wichtige *Monitoring- und Benchmark-Informationen* zur Qualität des Schulsystems an die Hand zu geben und

... dabei zu helfen, ein Verständnis für mögliche Erklärungen der festgestellten Differenzen zwischen Bildungssystemen zu entwickeln und entsprechende Diskussionsgrundlagen zu liefern sowie Ansatzpunkte für *steuernde Maßnahmen* zu erkennen (vgl. z.B. Plomp 2001).

Tabelle 1 gibt eine Übersicht über die *IEA-Studien mit deutscher Beteiligung*, zu der sich mit PISA auch eine OECD-Studie gesellt hat. Die erste Hauptuntersuchung stellt die *'First International Mathematics Study' (FIMS)* mit ihrer Erhebung in 1963/64 in 12 Ländern mit insgesamt 133.000 SchülerInnen der Sekundarstufen I und II dar (vgl. Husén 1967): „[...] pupils should be sampled at two points that could be regarded as ‚strategic‘ in the school system, that is toward the end of compulsory schooling and at the end of the preuniversity (academic secondary) school“ (ebd., S. 33). Das Prinzip der Testung an zentralen Punkten der Schullaufbahn ist bis heute bestehen geblieben. Zudem erfahren mathematisch-naturwissenschaftliche Kompetenzen immer noch besondere Aufmerksamkeit und Wertschätzung. Gleichwohl wird im FIMS-Report vermerkt, dass Mathematik selbstverständlich nicht der einzige Index der Produktivität eines Systems sei, sich das Fach aber wegen der – im Vergleich zu anderen Inhalten – methodisch einfacheren Handhabung für eine erste Erhebung dieses Umfangs besonders eigne (Husén 1967, S. 33-34). Zudem gab es internationale eine intensive Diskussion über die Gestaltung mathematischer Curricula und entsprechender Lehrmethoden: „[...] a movement of reconstruction of mathematical education“ (ebd., S. 77).

Die erste internationale Mathematikstudie wurde – wie heute noch bei vergleichbaren Erhebungen üblich – durch Hintergrundbefragungen von Schüle-

rInnen, Lehrkräften und Schulleitung zu schulischen und außerschulischen Merkmalen ergänzt, eingebettet in den Kontext der nationalstaatlichen Kultur, Geschichte und der sozialen Strukturen.

An der zweiten *Mathematikstudie (SIMS)* in den frühen 80er Jahren nahm Deutschland nicht teil, erst die dritte IEA-Untersuchung zu mathematischen kombiniert mit naturwissenschaftlichen Leistungen Mitte der 1990er Jahre fand mit deutschen SchülerInnen statt und verhalf empirischen Vergleichsuntersuchungen unter dem Akronym TIMSS zu großer Aufmerksamkeit (vgl. <http://timss.bc.edu/>, Stand: 2/02 und Baumert u.a. 1997, Baumert/Bos/Watermann 1998 sowie Baumert/Bos/Lehmann 2000).

Im Bereich der Testung naturwissenschaftlicher Inhalte wurden ebenfalls große IEA-Leistungsmessungen durchgeführt, bei denen alle westdeutschen Bundesländer wiederum nur in die erste und die letzte – ‚*First International Science Study*‘ FISS im Jahr 1970 (vgl. Comber/Keeves 1973) und TIMSS – involviert waren. Die ‚*Second International Science Study*‘ zwischen 1983 und 1986 wurde aus deutscher Sicht ausgelassen. Eine mögliche Erklärung für die *weitgehende Abstinenz Deutschlands während der 70er und 80er Jahre* – nach der „halbherzigen Teilnahme“ an der ersten Mathematikstudie mit nur zwei Bundesländern (Hessen und Schleswig-Holstein) – sehen Bos und Postlethwaite in der „eher geisteswissenschaftlichen Tradition deutscher Pädagogik“ (2001, S. 253); die systematische empirische Schulleistungsforschung hat erst seit dem „TIMSS-Schock“ einen anderen Stellenwert erhalten.

Eine Erklärung der deutschen Zurückhaltung in Bezug auf eine Beteiligung an Large Scale Assessments könnte in diesem Kontext auch die Frage nach der vernünftigen Relation zwischen den Ergebnissen derartiger Studien, deren Interpretation damals noch mit vielfältigen methodischen Problemen behaftet war, und den anfallenden Kosten gewesen sein: „Wie vielfach bei solchen Untersuchungen wird hierbei ein Aufwand bei der statistisch rechnerischen Verarbeitung der Daten betrieben, der nicht immer in einem angemessenen Verhältnis zu einigen grundlegenden Fakten, die notwendig mit solchen Untersuchungen verbunden sind, zu stehen scheint“, schreiben Schultze und Riemenschneider im deutschen FIMS-Bericht (1967, S. 32). Weiterhin heißt es: „Infolge dieser Unbestimmtheit bzw. Unklarheit der Kriterien und der Fragestellungen, die durch die Übersetzung in seltene Sprachen (Hebräisch, Japanisch) noch verstärkt wird, ist ein Großteil der Ergebnisse vage und nicht ausdeutbar“ (ebd., S. 33).

Ein Blick auf andere Kompetenzbereiche wie das *Leseverständnis* spiegelt eine ähnliche Zurückhaltung wider: Die großflächige, internationale und empiriegestützte Evaluation ist in Deutschland erst seit den 90er Jahren ein Thema intensiver erziehungswissenschaftlicher Forschung und bildungspolitischer Aufmerksamkeit. Im Bereich des Lesens sind hier vor allem die ‚*International Study of Reading Literacy*‘ (IRLS) von 1990/91 (vgl. Elley 1994 und Lehmann u.a. 1995), PISA mit dem aktuellen Schwerpunkt ‚Reading Literacy‘ (vgl. Deutsches PISA-Konsortium 2001 und OECD 2001) und die ‚*Internationale Grundschul-Lese-Untersuchung*‘ (IGLU) als nationale Bezeichnung der international angelegten Studie ‚*Progress in International Reading Literacy Study*‘ (PIRLS) zu nennen (vgl. www.erzwiss.uni-hamburg.de/IGLU/start.htm und <http://timss.bc.edu/pirls2001.html>, Stand: 2/02). Die Ergebnisse der letz-

teren Untersuchung werden für 2002/03 erwartet. Mittlerweile findet sich die Bundesrepublik auch bei Erhebungen etwa zum politischen Wissen und zu politischen Einstellungen (‚*Civic Education Study*‘, CivEd, vgl. http://www2.rz.hu-berlin.de/empir_bf/iea_e.html, Stand: 2/02 und Torney-Purta 2001) oder zu informationstechnologischem Wissen (‚*Computers in Education Study*‘, CompEd, vgl. Pelgrum/Plomp 1991 und 1993 sowie Pelgrum/Janssen Reinen/Plomp 1993) wieder.

2.2 Internationale Schulleistungsforschung aus deutscher Sicht

Die ersten IEA-Erhebungen waren mit einer *Vielzahl methodischer Probleme* verknüpft; so konnten nur wenige Bildungssysteme, die an den ersten Studien teilgenommen haben, die geforderte Ausschöpfungsquote von 85 % als Anteil der Personen der Bruttostichprobe, der tatsächlich getestet wurde, erreichen. Daher sind die Ergebnisse mit Vorsicht zu betrachten; sie sind weniger definitive Resultate als vielmehr Trendindikatoren. Die aktuelle PISA-Studie hingegen ist methodisch ausgereift. Zudem haben bis auf die Niederlande alle Teilnehmerstaaten die methodischen Anforderungen erfüllt.

Die Tabellen 2 und 3 vermitteln eine grobe Orientierung im Bereich der deutschen Leistungsentwicklung bei exemplarisch ausgewählten Large Scale Assessments im Vergleich zu den Ergebnissen anderer mitteleuropäischer Länder, im Kontrast zu nordeuropäischen Staaten und bezüglich ausgewählter anglophoner und asiatischer Überseegebiete. Die nationalen Leistungen wurden mittels des Rangplatzes bzw. der mittleren erreichten Punktzahl einem *oberen, mittleren bzw. unterem Leistungsfeld* zugeordnet. So wird zum einen sichtbar, welche Länder bei welchen Studien dabei waren, zum anderen lassen sich Tendenzen der Leistungsentwicklung zwischen den Ländern auf Anhieb erkennen. Die Reihenfolge der Studien ist auf den Erhebungszeitraum bezogen.

Die 1967 bzw. 1973 publizierten Resultate der FIM- und FIS-Studien mussten wie zur Bestätigung des von den OECD-Experten 1971 konstatierten *Moderitätsrückstandes* (s.o.) wirken:

Die Nicht-Mathematik-SchülerInnen¹ erreichten bei FIMS die absolute Spitzenposition in einem Feld von 10 Ländern vor Frankreich und Japan. Allerdings ist die Zuordnung der Schülergruppen zur mathematischen Grundbildung und zur voruniversitären Ausbildung für die anderen Länder mit ihren spezifischen Strukturen nicht eindeutig. Möglicherweise haben sich hier Unterschiede der Unterrichtszeit und -organisation ausgewirkt. So schreiben Schultze und Riemenschneider 1967 in Bezug auf die Vergleichbarkeit der englischen und der deutschen Resultate: „In der 6th Form der Grammarschool wer-

¹ Die erste IEA-Mathematikstudie (FISS) unterscheidet ‚Math-Science‘- bzw. ‚Nonmath-Science-Students‘. Grundsätzlich werden damit solche Schülergruppen getrennt voneinander betrachtet, die entweder Mathematik bzw. Naturwissenschaften als wesentlichen Bestandteil ihrer voruniversitären Ausbildung zum damaligen Zeitpunkt oder Mathematik/naturwissenschaftliche Fächer eher in Grundbildungsgängen ergänzend zu anderen Kernfächern belegt haben. Bezogen auf Deutschland wurden die SchülerInnen der mathematisch-naturwissenschaftlichen Gymnasien ersterer Population zugeordnet, die SchülerInnen der alt- und neusprachlichen Gymnasien fielen in die zweite Kategorie.

Tabelle 2: (Isabell van Ackeren): Leistungsentwicklung ausgewählter Bildungssysteme bei internationalen Large Scale Assessments seit den 1960er Jahren – Mittel- und nordeuropäische Staaten – Die grobe qualitative Einordnung der Rangplätze (oberes ↑, mittleres =, unteres ↓ Leistungsfeld) soll eine erste Übersicht ermöglichen. Neben dem Rangplatz im Verhältnis zur Anzahl der insgesamt beteiligten Systeme wurde auch die mittlere Punktzahl herangezogen.

Studie	FIMS			FISS		IRLS		TIMSS			CIVED			PISA		
	Math	Math	Nicht-Math	Mat.wiss.	Lesen	NaW	Ma	NaW	Ma	NaW	Ma	NaW	Ma	NaW	Ma	NaW
Inhalte	13 Jahre	12 Jahre	10 Jahre	10 Jahre	9 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	15 Jahre	15 Jahre	15 Jahre
Altersgruppe/Stufe	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII
Anzahl beteiligter Bildungssysteme	12	12	10	14	14	14	14	14	41	41	21	21	21	31	31	31
Deutschland	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
England	5	7	1	10	5	2	16 NW 18 O 16 W	13	18	23	12	13	15	21	20	20
Frankreich	6	↑	7	8	9	6	↑	17	↑	↑	↑	↑	17	7	8	4
Niederlande	9	↑	2	↑	10	↑	↑	4	28	13	7	7	14	10	12	12
Österreich	8	↑	4	9	13	4	↓	21	↑	9	2	1	10	11	8	8
Schweiz	↑	↑	5	↑	↑	↑	↑	↑	6	6	↑	↑	10	11	↑	8
Dänemark	↑	↑	↑	↑	↑	↑	↑	↑	4	8	↑	↑	10	11	↑	8
Finnland	↑	↑	↑	↑	↑	↑	↑	↑	25	8	↑	↑	19	17	7	18
Norwegen	↑	↑	↑	↑	↑	↑	↑	↑	34	27	↑	↑	14	16	12	22
Schweden	↓	↓	↓	↓	↓	↓	↓	↓	17	19	↓	↓	9	13	17	13
	12	8	9	2	6	9	3	3	16	22	1	2	18	9	15	10

Tabelle 3: (Isabell van Ackeren): Leistungsentwicklung ausgewählter Bildungssysteme bei internationalen Large Scale Assessments seit den 1960er Jahren – Englischsprachige Überseegebiete und asiatische Staaten – Die grobe qualitative Einordnung der Rangplätze (oberes ↑, mittleres =, unteres ↓ Leistungsfeld) soll eine erste Übersicht ermöglichen. Neben dem Rangplatz im Verhältnis zur Anzahl der insgesamt beteiligten Systeme wurde auch die mittlere Punktzahl herangezogen.

Studie	FIMS			FISS		IRLS		TIMSS			CIVED			PISA		
	Math	Math	Nicht-Math	Mat.wiss.	Lesen	NaW	Ma	NaW	Ma	NaW	Ma	NaW	Ma	NaW	Ma	NaW
Inhalte	13 Jahre	12 Jahre	10 Jahre	10 Jahre	9 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	14 Jahre	15 Jahre	15 Jahre	15 Jahre	15 Jahre
Altersgruppe/Stufe	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII	SII
Anzahl beteiligter Bildungssysteme	12	12	10	14	14	14	14	41	41	21	21	21	31	31	31	31
Australien	↓	↓	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
England	10	11	3	3	3	5	11	12	16	9	11	11	4	5	7	7
Kanada	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
Neuseeland	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
USA	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
	11	12	4	7	14	2	9	17	19	5	9	9	2	6	5	5
Hongkong	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
Japan	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
Korea	2	6	1	1	1	2	3	3	3	3	3	3	8	1	2	1
Singapur	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
	11	12	10	4	7	14	2	9	17	28	16	19	15	19	14	14

den von vielen Schülern nur drei Prüfungsfächer betrieben; infolge dieser weitgehenden Spezialisierung können bei den Mathematikern relativ hohe Leistungen erwartet werden, während in der BRD die Spezialisierung in Bezug auf Mathematik nur gering ist, da Mathematik in jedem Fall als Abiturfach auftritt“ (Schultze/Riemenschneider 1967, S. 15). Die internationale Vergleichbarkeit der Schülerpopulationen bedürfte folglich eines genaueren Blicks.

Im Bereich der mathematischen Grundbildung schien – einmal abgesehen von möglichen methodischen Problemen – das deutsche Schulsystem für die höhere Schulbildung zu funktionieren (vgl. Tabelle 2); auf dem Gebiet der mathematischen Spitzenleistungen damals – wie heute bei PISA – allerdings schon nicht mehr: Die deutschen ‚mathematics students‘ erreichten gegenüber ihren internationalen KollegInnen nur einen Platz im Mittelfeld: „[...] die Leistungen unserer Abiturienten im mathematisch-naturwissenschaftlichen Zweig sind bescheiden gegenüber den Leistungen, die in anderen Ländern von sehr viel jüngeren Schülern zur Zeit ihres Übergangs an die Universität erzielt werden“ (Schultze/Riemenschneider 1967, S. 34). Auch die deutschen 13-Jährigen aus Volks-, Realschulen und Gymnasien schnitten mit Rang 5 von insgesamt 12 Systemen nur noch mittelmäßig ab hinter einer Spitzengruppe bestehend aus Israel, Japan und Belgien.

Die FISS-Ergebnisse belegten auch im naturwissenschaftlichen Bereich ein *bedenkliches ‚Abrutschen‘* von guten Leistungen der älteren SchülerInnen bis hin zu mittelmäßigen bzw. unterdurchschnittlichen Ergebnissen unterer Jahrgangsstufen: Während die älteren deutschen SchülerInnen (aus den Abschlussklassen der Sekundarschulen) in einem Feld von 14 Ländern – darunter Australien, England, Finnland, Schweden und die USA – noch nach Neuseeland den zweiten Rang erreichten, lagen die 14-Jährigen ‚nur‘ noch im oberen Mittelfeld auf Rang 5 – hinter Japan (damals schon auf dem Spitzenplatz), Ungarn, Australien und Neuseeland. Die 10-Jährigen dagegen lagen auf dem drittletzten Rang: Japan, Schweden, Finnland, die Niederlande und andere lagen – z.T. mit deutlichem Abstand – vor den SchülerInnen aus der BRD.

Unter dem – offensichtlich zeitlosen – Titel „Im internationalen Vergleich schneidet das Bildungswesen der Bundesrepublik miserabel ab: Die deutschen Schüler auf dem letzten Platz“ schrieb Hayo Matthiesen am 20.9.1974 mit Blick auf diese Ergebnisse von FISS: „Vor allem gegen die Effektivität des von vielen so gelobten traditionellen dreigliedrigen Schulwesens liefert die Studie Belege. Rigoros wie sonst kaum irgendwo in der Welt liest dieses System die Schüler aus, so dass sich bei uns in den Abschlussklassen der Sekundarstufe II nur neun Prozent eines Jahrgangs befinden; in den USA sind es 75 %, 70 in Japan, 47 in Belgien oder 29 in Australien [...]“.

Ungeachtet der insgesamt nicht überzeugenden deutschen Leistungen bei diesen ersten Vergleichsuntersuchungen beteiligte sich die BRD *nach Jahren der Abstinenz* erst zu *Beginn der 90er Jahre* mit der internationalen *Lesestudie IRLS* wieder an einem solchen Projekt. In der jüngeren IRLS-Zielgruppe (Drittklässler) und der älteren Altersgruppe (8. Jahrgangsstufe) ist das Leseverständnis in beiden Teilen Deutschlands (die Studie wurde vor der Wiedervereinigung begonnen) *insgesamt nur durchschnittlich* ausgeprägt; besonders stark fällt dies im Vergleich zu ökonomisch entwickelten Ländern auf. Ebenfalls ungünstig sehen die Ergebnisse im Bereich der Leistungsstreuung aus (vgl. Leh-

mann u.a. 1995, S. 217). Es scheint folglich auch am Beginn der Schulkarrieren nicht zu gelingen, einen guten Mindeststandard für die Altersgruppe insgesamt zu sichern. Zumindest gilt dies für das Leseverständnis; an der TIMSS-Grundschuluntersuchung zu Mathematik und Naturwissenschaften war Deutschland nicht beteiligt, so dass für diese Altersgruppe in diesen Fächern keine Daten vorliegen. Umso spannender dürften die Ergebnisse der *internationalen Grundschul-Lese-Untersuchung* sein (IGLU/PIRLS), die für Ende 2002/Anfang 2003 erwartet werden. Die Ergebnisse der IRL-Studie wurden erst rückwirkend nach der TIMSS-Rezeption Mitte der 90er Jahre von der Fachöffentlichkeit wahrgenommen; vermutlich erschienen mathematisch-naturwissenschaftliche Kompetenzen in der ökonomisch orientierten Standortdiskussion wichtiger als Leseverständnis und -gewohnheiten. Mittlerweile scheint das Bewusstsein gewachsen zu sein, dass Leseverständnis eine grundlegende Voraussetzung mathematischer Kompetenz darstellt, wie es durch PISA einmal mehr bestätigt wurde.

Die nur mittelmäßigen Ergebnisse der *TIMS-Studie* dürften hinlänglich bekannt sein. Für Deutschland waren die Befunde insgesamt – für alle Bildungsgänge, für die unterschiedlichen Kompetenzbereiche und für alle Bundesländer – sehr mittelmäßig. Eine wesentliche Wirkung der Erhebung war die Zäsur, die durch TIMSS im deutschen Umgang mit Large Scale Assessments erreicht wurde. Dieser Trend der Mittelmäßigkeit hat sich bei PISA bestätigt bzw. es deutet sich – betrachtet man die erreichten Rangplätze – eher ein *weiterer Abwärtstrend* an.

Die wie bei PISA im Jahr 2001 veröffentlichten Resultate der „*Civic Education Study*“ zu *Demokratie und Bildung* in 28 Ländern hat bisher im Gegensatz zu PISA kaum Aufmerksamkeit erfahren. Dies mag auch damit zusammenhängen, dass bis dato nur ein internationaler Bericht vorliegt (Torney-Purta u.a. 2001), der um länderspezifische Berichte zum politischen Verstehen und Engagement bei 14-Jährigen ergänzt werden soll. Der Bericht, der auch in einer deutschen Zusammenfassung im Internet abrufbar ist (www2.hu-berlin.de/empir_bf/reports.html, Stand: 1/2002), bedarf einer tiefer gehenden Rezeption. Für die Übersicht in Tabelle 2 wurde der Aspekt ‚politisches Wissen/Verstehen‘ herausgegriffen. Die deutschen SchülerInnen nehmen auch in diesem Kompetenzbereich einen *mittleren Rang* ein, der dem internationalen Mittelwert entspricht.

Die Tabellen 2 und 3 ermöglichen nicht nur eine Betrachtung der Leistungsentwicklung bei Vergleichsstudien aus deutscher Sicht. Darüber hinaus wurden Staaten in den Überblick einbezogen, die *wichtige Vergleichsländer* darstellen: Nachbarländer, wirtschaftlich ähnlich entwickelte Nationen, Staaten mit vergleichbaren kulturellen Traditionen und solche, die bei PISA sehr gut abgeschnitten haben. Einige dieser Länder werden nachfolgend in Ergänzung zur Übersicht herausgegriffen, in ihrer Leistungsentwicklung dargestellt und mit möglichen erklärenden Hintergrundmerkmalen, die sich auf Reformen der letzten Jahre beziehen, verknüpft. In diesem Rahmen können dies nur aller erste, fragmentarische Hinweise sein, denen man im Einzelnen nachgehen sollte. Hier kommt der Vergleichenden Erziehungswissenschaft in Verknüpfung mit der Schulleistungsforschung und der Schulentwicklungsforschung eine besondere Bedeutung zu.

2.3 Leistungen in anderen Bildungssystemen

2.3.1 Finnland

Dem Betrachter der Ranglisten zu den aktuellen PISA-Ergebnissen fällt zunächst die *Vorreiterstellung Finnlands* in allen Testbereichen mit einem 1. Rang bei der Lesekompetenz, einem 4. Platz in Mathematik und einer Positionierung an 3. Stelle im Bereich der Naturwissenschaften ins Auge. Nicht nur der Blick auf die erreichten Ländermittelwerte zeigt, dass die finnischen SchülerInnen einen Platz im absoluten Spitzenfeld erreichen konnten. Es gelingt dem Bildungssystem überdies, bei einem hohen Durchschnittswert die gemessenen Einzelwerte vergleichsweise dicht um diesen Mittelwert zu positionieren, d.h. die *Leistungsspreizung fällt gering aus* mit einer äußerst kleinen Risikogruppe am unteren und ausgeprägten Spitzenleistungen am oberen Ende der Skala.

Dieses noch deutlich bessere Abschneiden der finnischen gegenüber den bereits überdurchschnittlich guten Leistungen der schwedischen Jugendlichen war sicherlich überraschend. Das *schwedische Schulsystem* ist mit dem ihm oft zugeschriebenen Modellcharakter in Verbindung mit seinem hervorragenden Abschneiden bei der TIMS-Oberstufenstudie im Bereich internationaler Vergleichsstudien in den Köpfen der Öffentlichkeit weit aus mehr präsent als Finnland. Die finnischen SchülerInnen waren bei keiner der TIMS-Untersuchungen beteiligt; bei der vorhergehenden Lesestudie zu Beginn der 90er Jahre waren sie aber bereits in beiden Altersgruppen mit Platz 1 absoluter Spitzenreiter; man befand sich damals zwei Plätze vor Schweden. Ein Blick auf die ersten Leistungsstudien FIMS und FISS zeigt sowohl für Schweden als auch für Finnland ein weniger stabiles Leistungsbild (vgl. Tabelle 2). Gerade seit den 70er Jahren hat es aber in vielen europäischen Ländern tief greifende Reformen gegeben, deren Effekte sich möglicherweise in den Schülerleistungen im Zeitverlauf widerspiegeln.

Innerhalb der Gruppe der nordeuropäischen Länder steht Finnland nun seit dem 4. Dezember 2001 im Zentrum des (fach)öffentlichen Interesses. Man geht und wird in Zukunft verstärkt der Frage nachgehen, welche Strukturen das finnische Bildungswesen kennzeichnen und in welchem möglichen Kausalzusammenhang sie zu den erreichten Leistungen stehen. Die Klärung dieses Zusammenhangs bedarf weiterer Studien. Dennoch geben die *strukturellen Merkmale des finnischen Systems* Aufschluss über eine in vielen Aspekten vom deutschen Bildungswesen abweichende *Möglichkeit der institutionellen Gestaltung von Bildungsprozessen*, die sichtbar – sicherlich in einer nationalspezifischen Kombination mit weiteren erklärenden Faktoren – zu Erfolgen führt.

Folgt man der Chronologie der Schullaufbahn, so haben in Finnland alle Eltern einen Anspruch auf einen Platz für ihre Kinder im Elementarbereich. Im Anschluss an den Vorschulunterricht, dem ein eigener Bildungsplan zugrunde liegt, werden die Kinder mit 7 Jahren in die gemeinsame Grundschule eingeschult. Finnland verfügt seit 1970 – nach einer kontinuierlichen Umstellung von einem *ehemals gegliederten zu einem integrierten Schulsystem* – über eine neunjährige Grundschule. Erst im Sekundarbereich II differieren die Schularbeiten und die entsprechenden Abschlüsse. Die finnischen Kinder und Jugendlichen lernen während ihrer gesamten Schullaufbahn in *Ganztagseinrichtungen*, was international eher den Normalfall darstellt. Im Bereich der Leis-

tungsmessung und -bewertung entscheiden die Schulen eigenständig über ihre Verfahren. Erst am Ende der Sekundarstufe I greifen *Mechanismen einer externen Wirkungssteuerung*: Dann nämlich gibt es eine *verpflichtende Abschlussprüfung*, die am Ende der Sekundarstufe II landesweit einheitlich zur gleichen Zeit unter Aufsicht einer vom Minister eingesetzten Kommission stattfindet, der u.a. Hochschulprofessoren angehören, um die abgebenden mit den aufnehmenden Institutionen stärker miteinander zu verzahnen. Diese ‚matri-culation examination‘ hat eine über 100-jährige Tradition, an der die SchülerInnen wahlweise zu drei Terminen im Jahr teilnehmen können (vgl. Vuorio-Lehti 1998). Verpflichtend in die nationale Prüfung zu nehmen sind die Muttersprache, die zweite Landessprache, eine Fremdsprache, Mathematik oder Naturwissenschaften oder auch Gesellschaftswissenschaften; es besteht die Option, sich in weiteren Fächern der nationalen Testung zu unterziehen.

Die finnischen Zeitungen erstellen *regelmäßig Ranglisten* auf der Grundlage der Prüfungsergebnisse (Kirjavainen/Loikkanen 1998). Auch das Ministerium kümmert sich um die Ergebnisse und meldet den Schulen regelmäßig die Leistungsergebnisse im nationalen Vergleich zurück. Da seit 1994 vor allem in größeren Städten eine *freie Schulwahlmöglichkeit* unabhängig von Einzugsgebieten besteht und sich Eltern bei ihrer Entscheidung für oder gegen eine bestimmte Schule auf die nationalen Rankings stützen können, hat sich in Finnland – wie in vielen anderen europäischen Ländern – eine *marktähnliche Situation* herausgebildet (Hirvenoja 1999). Um den Schulen die Möglichkeit flexibler Reaktionen auf die Bedingungen dieses Marktes durch einzelschulische Profilbildungen zu geben, wurde ebenfalls 1994, 20 Jahre nach seiner Implementierung, das traditionelle präskriptive Nationale Curriculum abgeschafft, das auf ein *Kerncurriculum* mit Kerninhalten für die Fächer, die Festlegung für die Prüfungskriterien und die Stundentafeln reduziert wurde.

Damit ergibt sich für Finnland auf der einen Seite ein Bild einer in den letzten Jahren *stark erweiterten schulischen Autonomie* im Bereich der Kontext- und Prozesssteuerung, auf der anderen Seite findet sich als funktionales Gegenstück und Möglichkeit der Qualitätssicherung und Vergleichbarkeit das *Instrument zentraler Prüfungen*, die wahrscheinlich in nicht unerheblichem Maße auf die den Prüfungen vorgelagerten Unterrichtsprozesse rückwirken. Mögliche Effekte der Lockerungen der Strukturen auf der Input- und Prozessseite (Stichwort reduziertes ‚Kerncurriculum‘ und eigenständige Festlegung der ‚laufbahnbegleitenden Leistungsmessung und -bewertung‘) zugunsten weiterer schulischer Gestaltungsprozesse sind im Vergleich der TIMSS- zu den PISA-Ergebnissen nicht sichtbar.

2.3.2 Die Niederlande

Das Prinzip relativ großer *Gestaltungsfreiheit und nationaler Standardsetzung* durch zentrale Abschlussprüfungen findet sich beispielsweise ebenso in den Niederlanden wieder (vgl. z.B. van Ackeren/Klemm 2000). Auch die Niederlande haben bei den Vergleichsstudien der 1990er Jahre durchweg Spitzenpositionen eingenommen. Bei PISA wäre es ein 1. Rang in Mathematik, ein 3. im Leseverständnis und ein 6. Platz in den Naturwissenschaften gewesen. Da allerdings die geforderten Ausschöpfungsquoten eines Altersjahrgangs nicht erreicht wurden und vermutet werden könnte, dass sich nur gute Schulen beteiligt haben, sind die Ergebnisse sehr kritisch zu betrachten.

Abgesehen von solchen Unsicherheitsfaktoren machen die Prüfungsstrukturen der überwiegenden Mehrheit der besten zehn PISA-Teilnehmerländer deutlich, dass dort *zentralisierte Prüfungssysteme eine große Rolle* spielen. Zudem kennt man in vielen dieser Länder auch *diagnostische Tests während der Schullaufbahn*, die nicht unbedingt an Gelenkstellen der Schulkarrieren stattfinden, sondern so positioniert sind, dass sie lediglich die *diagnostische Kompetenz der Lehrkräfte* im nationalen Vergleich unterstützen sollen; eine Qualifikation, die offenbar bei den deutschen Lehrkräften nicht hinreichend ausgeprägt ist, wie die PISA-Studie andeutet. Der Standardsicherung durch externe Evaluationsmaßnahmen im unterstützenden und weniger kontrollierenden Sinne nachzugehen, scheint sich zu lohnen.

2.3.3 England

Ein erhöhter Aufmerksamkeitsbedarf für die Testung von Mindeststandards wird vielleicht auch durch das Abschneiden Englands bei PISA bestätigt: Innerhalb der Gruppe der mitteleuropäischen Nachbarländer Deutschlands fällt für England die *enorme Leistungssteigerung in Mathematik bei PISA* (hier war das Vereinigte Königreich insgesamt beteiligt) im Vergleich zum sehr mittelmäßigen Ergebnis (noch hinter Deutschland) bei der TIMSS-Primar- und -Mittelstufenstudie auf; an der Oberstufen-Untersuchung war England nicht beteiligt, ebenso nicht an der Lesestudie von 1990/91, so dass in diesem Kompetenzbereich keine Ergebnisse verglichen werden können. Die naturwissenschaftlichen Ergebnisse waren hingegen bei TIMSS bereits überzeugend; bei PISA wurde nun aber der absolute Spitzenbereich erreicht. Auch wenn viele englische Zeitungen um den 4. Dezember 2001, dem Tag der PISA-Veröffentlichung, mit Headlines wie „British children near top of world class“ (O’Leary/Owen 2001) über das englische Abschneiden bei PISA berichten, haben sie ebenso schnell den weniger erfreulichen Aspekt aus britischer Sicht im Blickfeld ihrer Artikel: „[...] research [...] showed it has one of the worst records of industrialised nation at getting pupils from poor families to do well at school“ (ebd.). Dies ist allerdings kein neuer Befund für das Vereinigte Königreich: Es gibt einen traditionellen Zusammenhang zwischen ‚social class‘ und ‚educational achievement‘. Dennoch bleibt das insgesamt sehr gute und deutlich gegenüber Vorgängerstudien verbesserte englische Abschneiden erklärungsbedürftig. Bildungsministerin Estelle Morris führt dies auf die *Reformen der letzten Jahre* zurück (ebd.). Folgende Erklärungen werden in den ersten britischen Presseartikeln zu PISA von zitierten Erziehungswissenschaftlern gegeben (vgl. u.a. Slater 2001):

Im Unterschied zu TIMSS war das Untersuchungsdesign von PISA auf ein Abtesten von *Wissen und Fertigkeiten in „real-world scenarios“* angelegt. Dies kommt den Inhalten des englischen Nationalen Curriculums sehr entgegen, bei denen in den letzten 15 Jahren, seit der großen Bildungsreform von 1988, besonderer Wert auf Problemlösefähigkeit im realitätsnahen Kontext gelegt wird. Zudem ist der naturwissenschaftliche Unterricht fächerübergreifend als Unterrichtsfach ‚Science‘ angelegt. Hier und in vielen anderen Ländern sind *Naturwissenschaften als Hauptfach implementiert* und oftmals nicht in einzelne Fächer aufgliedert, die nicht zusammenhängend über die Schullaufbahn hinweg unterrichtet werden.

Die verbesserten Ergebnisse werden auch auf die Bemühungen um eine *Leistungssteigerung von schwachen Lernern* zurückgeführt. Auf diese Lerngruppe

hat man sich in den letzten Jahren besonders konzentriert und dies wird vom deutschen PISA-Konsortium auch für die Bundesrepublik empfohlen, um das Leistungsniveau insgesamt zu heben.

Auf der anderen Seite wird auch vermerkt, dass die Resultate möglicherweise inflationär seien, da die Ausschöpfungsquote vergleichsweise gering war: „Maybe fewer low-achieving schools in the UK took part.“

Möglicherweise spielt die Ministerin – um auf die externen Qualitätskontrollen als standardsichernde Maßnahmen zurückzukommen – mit ihrem Verweis auf die Bildungsreformen der letzten Jahre als Motor verbesserter schulischer Leistungsergebnisse auch auf die 1988 eingeführten *regelmäßigen Assessments an vier Stellen der Schullaufbahn* in Ergänzung zu den *traditionellen Abschlussprüfungen* an. Es ist nicht auszuschließen, dass diese und andere Maßnahmen – etwa die Einführung eines Nationalen Curriculums – bei TIMSS noch nicht durchgeschlagen haben, aber mittlerweile bis zu PISA durchgewachsen sind. Ähnliche Argumente finden sich im Übrigen auch im Hinblick auf die Leistungssteigerung der USA, bei denen ein Abrutschen von sehr guten Resultaten bei der TIMSS-Grundschuluntersuchung über mittelmäßige Ergebnisse in der Sekundarstufe I bis hin zu unteren Rangplätzen im Sekundarbereich II zu verzeichnen war. Vermutlich sind bei der signifikanten Leistungssteigerung der USA bei PISA die Effekte standardsichernder, externer Kontrollmechanismen, die bei der Grundschulstudie bereits sichtbar waren, nun auch in den älteren Jahrgängen erkennbar.

2.3.4 Andere Länder

Neben diesen regelmäßigen Studien kennen viele der leistungsstarken Länder auch *lokale Studien*, die den deutschen bundesländerinternen Untersuchungen wie LAU, QuaSUM oder MARKUS in ihrer grundsätzlichen Anlage vergleichbar sind (vgl. van Ackeren 2002). Dies ist etwa in Frankreich und den Niederlanden der Fall und macht Schulen zu datenreichen Institutionen, die sie möglicherweise auch für Schulentwicklungsprozesse nutzen. Ein ähnliches Projekt ist ebenso auf der Homepage des finnischen Bildungsministeriums beschrieben: Seit 1996 gibt es dort ein spezielles Förderprogramm für Mathematik- und Naturwissenschaften, in dem die Leistungen der finnischen SchülerInnen evaluiert werden. Darüber hinaus werden auch gezielt die Curricula evaluiert und weiterentwickelt, ebenso Unterrichtsmaterialien, Lehr- und Lernmethoden und Lehrerbildungsprogramme. Es wurde ein *nationales Netzwerk* aufgebaut, an dem auch WirtschaftsvertreterInnen beteiligt sind. Man konzentriert sich folglich in Finnland auf das gesamte Spektrum der institutionellen Bedingungen schulischen Lernens und steht darüber offenbar in einem *breiten öffentlichen und vor allem strukturierten Diskurs*.

Auch die im Zeitverlauf gesehene Leistungssteigerung und -stabilisierung *Australiens, Kanadas und Neuseelands* ist sicher eine weitere Analyse wert. Hilfreich sind hier möglicherweise auch die einzelnen PISA-Länderberichte mit möglichen Erklärungen zu entsprechenden landesspezifischen Ursache-Wirkungs-Zusammenhängen. Auch hier scheinen regelmäßige Monitoring-Prozesse eine wichtige Rolle einzunehmen, wie etwa für Neuseeland beschrieben: „Assessment and reporting procedures are designed to improve a rich picture of what children can do [...]. The result is a detailed national picture

of student achievement“ (Crooks 2001, S. 1). Die SchülerInnen in Neuseeland werden im Alter von 9, 12, 15 und 17 Jahren landesweit getestet; dies ist ein ähnlich dichtes Testnetzwerk wie in England. Zudem gibt es in Neuseeland eine so genannte Literacy bzw. Numeracy Strategy, wie man sie auch aus England kennt (vgl. Limbrick 1999). Im Zentrum steht in beiden Ländern, Leistungsprofile für die unterschiedlichen Alterstufen zu entwickeln und sich auf bestimmte Mindeststandards zu einigen. Im Fokus steht weiterhin die Entwicklung und Verbreitung effektiver Lehrmethoden und Diagnoseinstrumente. Viele der Maßnahmen beziehen sich vor allem auf die unterste Leistungsgruppe.

3. Fazit und Ausblick

Heute – nach TIMSS und PISA – können wir im Rückblick auf die seither vergangenen 40 Jahre feststellen: Das „Mangelhaft“ der OECD-Experten konnte ebenso wenig wie die Vorbote der absinkenden Konkurrenzfähigkeit deutscher Schülerinnen und Schüler verhindern, dass die westdeutsche BRD und das nach 1989 vereinigte Deutschland insgesamt den *Anschluss an den Modernisierungszug der entwickelten Länder verpasst* hat. Im Zentrum des vorangegangenen Überblicks über die Leistungsentwicklung in wichtigen Vergleichsstaaten standen Mechanismen der Wirkungssteuerung, zu der auch Studien wie PISA selbst gehören. Es sind zudem institutionelle Strukturen, die hierzulande – auch in der aktuellen PISA-Diskussion – weniger stark diskutiert werden. Die Perspektive darf selbstverständlich im Hinblick auf Hinweise für politische Handlungsmöglichkeiten nicht auf diesen Aspekt verengt werden. Deshalb soll der Blick auf die aktuellen PISA-Ergebnisse und mögliche Einflussfaktoren zum Schluss des Beitrags in internationaler Hinsicht kurz erweitert werden:

- Nur in Österreich und in zwei Kantonen der Schweiz werden Kinder schon im Alter von zehn Jahren auf unterschiedliche Bildungswege verteilt.
- Deutschland erzielt mit dieser frühen Auslese einen *Anteil der Studienberechtigten* am Altersjahrgang, der mit 33 % dramatisch unterhalb der entsprechenden Quoten anderer Länder liegt: Im Durchschnitt aller OECD-Länder erreichen 57 % eine Studienberechtigung, in Finnland – einem der Spitzenreiter bei der PISA-Studie – sind es sogar 89 %.
- Deutschland zählt zu den wenigen Ländern, die ihre SchülerInnen nur am Vormittag in den Schulen aufnehmen. Innerhalb der EU-Staaten haben nur Griechenland und Österreich die Halbtagsschule, Italien und Portugal bieten Mischformen ganz- und halbtägiger Betreuung, alle anderen Länder kennen die *Ganztagsschule als Regelschule*.
- Beim *Anteil öffentlicher Bildungsausgaben* am Bruttoinlandsprodukt liegt Deutschland mit 4,4 % deutlich unterhalb des OECD-Durchschnittswertes (5,0 %) und weit hinter Ländern wie Finnland (5,8 %), Schweden (6,6 %) oder Norwegen und Dänemark (je 6,8 %).
- Mit seinem früh und hoch selektivem sowie unterfinanziertem Schulsystem erzielt *Deutschland* – dies zeigen die Vergleichsstudien der neunziger Jahre und eindringlicher denn je die PISA-Untersuchung – international gesehen *bestenfalls mittelmäßige Leistungen*: und zwar gleichermaßen in Mathematik, in Naturwissenschaften und beim Leseverständnis, gleichermaßen in den unterschiedlichen Altersgruppen und gleichermaßen in den unterschiedlichen Leistungsgruppen. Die deutschen Spitzengruppen laufen ihren internationalen Kol-

legen nahezu genauso hinterher wie die schwachen deutschen SchülerInnen den ihnen vergleichbaren Gruppen anderer Länder.

- Schließlich und nicht zuletzt: In keinem der in der PISA-Studie untersuchten Länder gestaltet sich der Zusammenhang zwischen *sozialer Herkunft und erreichter Schulleistung* so eng wie in Deutschland. Kein anderes Land verfehlt das demokratische Ziel der Chancengleichheit so fundamental wie Deutschland.

Das Resümee ist bitter: So wie Deutschland in das 20. Jahrhundert mit einer – im Vergleich zu den anderen entwickelten Ländern der Welt – Spitzenposition im Bereich des Bildungswesens eintrat, so startet es in das 21. Jahrhundert weit abgeschlagen. Den didaktischen und strukturellen Modernisierungsprozess, den viele der Vergleichsländer in der zweiten Hälfte des 20. Jahrhunderts vollzogen und auch finanziert haben, glaubte Deutschland nicht nötig zu haben. Das Erwachen und der Blick auf den *verfestigten Modernitätsrückstand* sind bitter.

Dabei ist nicht sicher, ob sich das Land zu einem wirklichen *Kurswechsel* aufraffen wird. Einige Beobachtungen müssen besorgt stimmen: Die Finanzminister haben im vergangenen Jahr – allerdings vor der Veröffentlichung der PISA-Studie – einstimmig ihren Sparwillen zu Protokoll gegeben. Nach 2005, wenn die Schülerzahlen im Land sinken werden, müssten – so ihr einhelliges Votum – Mittel aus dem Bildungsbereich abgezogen werden. Die Kultusminister haben sich nach Veröffentlichung der PISA-Studie tief besorgt gezeigt, sie haben auch angekündigt, das gesamte Bildungssystem auf den Prüfstand stellen zu wollen, sie erklären aber zugleich, dass die *Struktur des deutschen Schulsystems* mit seiner frühen Auslese nicht zur Disposition stehen könne. Letztlich unverbindlich verweisen sie auf die Qualität des Unterrichts und damit vorrangig auf die Arbeit der LehrerInnen und Lehrer. Deren Ausbildung zu reformieren haben aber, mit geringen Nuancierungen, alle Bundesländer während der neunziger Jahre versäumt. Das große *Umsteuern in Deutschlands Schulen* soll und wird sich wohl *in den bestehenden Strukturen*, mit den begrenzten Finanzmitteln und mit Lehrenden vollziehen, die das konstatierte Elend nicht haben verhindern können und die, wenn sie neu in den Schuldienst eintreten, so wie ihre Altvorderen ausgebildet wurden.

Aber es winkt Ablenkung vom unterdurchschnittlichen Abschneiden im internationalen Vergleich: Wenn im Laufe dieses Jahrs die *innerdeutsche PISA-Auswertung* – mit den unvermeidbaren Länderranglisten – vorgelegt wird, dann kann das *innenpolitische Hauen und Stechen* anheben. Der Streit um den besten innerdeutschen Schulweg, der sich jetzt schon im Vorgriff auf nicht bekannte Landesergebnisse in Debatten im Bundestag und in einzelnen Landesparlamenten ankündigt, wird für einige Zeit vergessen machen, dass es in Wirklichkeit ein Streit darüber ist, *wer unter den im internationalen Vergleich insgesamt rückständigen Bundesländern am wenigsten rückständig ist*. Ein Blick auf die TIMS-Studie belegt dies: Bei der Präsentation der Mittelstufenstudie wurden die Ergebnisse von zwei anonymisierten Bundesländern gegenübergestellt. Wenn man das bessere dieser beiden Länder, das Land A (Insider lieben erkennen, dass es sich dabei um Bayern handelt habe) als einzelnes Land in die internationale Rangliste aufnehmen würde, so läge dieses Land auf dem 17. Platz. Zur Erinnerung: Deutschland insgesamt lag auf Platz 23.

Literatur

- van Ackeren, Isabell; Klemm, Klaus 2000: Liberté oder Egalité – Zwischen Wettbewerb und Gleichheit: Wie andere europäische Länder mit den Ergebnissen von Leistungsstudien in Schulen umgehen. In: *Süddeutsche Zeitung*, 28.11.2000, Seite V2/12. Online unter: www.uni-essen.de/bfp (Stand: 2/02)
- van Ackeren, Isabell 2002: Zentrale Tests und Prüfungen im Dienste schulischer Entwicklung – Erfahrungen in Deutschland sowie Beobachtungen aus England, Frankreich und den Niederlanden. In: *Bildung und Erziehung* Heft 1/März 2002 (im Erscheinen)
- Baumert, Jürgen u.a. 1997: TIMSS – mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen: Leske + Budrich
- Baumert, Jürgen; Bos, Wilfried; Watermann, Rainer 1999: TIMSS III. Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Zusammenfassung deskriptiver Ergebnisse. Berlin: Max-Planck-Institut für Bildungsforschung, 1999
- Baumert, Jürgen; Bos, Wilfried; Lehmann, Rainer (Hrsg.) 2000: TIMSS III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Bd. 1-2. Opladen: Leske + Budrich, 2000
- Bos, Wilfried; Postlethwaite, Neville 2001: Internationale Schulleistungsforschung. Ihre Entwicklungen und Folgen für die deutsche Bildungslandschaft. In: Weinert, Franz E. (Hrsg.): *Leistungsmessungen in Schulen*. Weinheim und Basel: Beltz
- Comber, L. C.; Keeves John P. 1973: *Science Education in Nineteen Countries. An Empirical Study*. Uppsala: Almqvist & Wiksell
- Crooks, Terry 2001: *New Zealand's National Education Monitoring Project (NEMP): rich information from multiple task formats*. Otago. Online unter: www.leeds.ac.uk/educol/ (Stand: 2/02)
- Deutsches PISA-Konsortium (Hrsg.) 2001: *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich
- Elley, Warwick B. 1994: *The IEA Study of Reading Literacy. Achievement and Instruction in Thirty-Two School Systems*. Exeter: Pergamon
- Herrlitz, Hans-Georg; Hopf, Wulf; Titze, Hartmut 1993: *Deutsche Schulgeschichte von 1800 bis zur Gegenwart. Eine Einführung*. Weinheim: Juventa
- Hirvenoja, Piia 1999: *Education policy changes and school choice in Europe from the Scandinavian perspective*. Turku. Online unter: www.leeds.ac.uk/educol/ (Stand: 2/02)
- Husén, Torsten (Hrsg.) 1967: *International Study of Achievement in Mathematics. A Comparison of Twelve Countries*. Stockholm: Almqvist & Wiksell
- Kirjavainen, Tanja; Loikkanen, Heikki A. 1998: *Up Or Down? On Changes In Student Performance In Finnish Senior Secondary Schools During 1980-92*. Tampere und Helsinki. Online unter: www.leeds.ac.uk/educol/ (Stand: 2/02)
- Lehmann, Rainer H. u.a. 1995: *Leseverständnis und Lesegewohnheiten deutscher Schüler und Schülerinnen*. Weinheim und Basel: Beltz
- Limbrick, Libby 1999: *The Literacy Debate: what are the issues in New Zealand*. Auckland. Online unter: www.leeds.ac.uk/educol/ (Stand: 2/02)
- Luehrig, Holger H. 1973: *Wirtschaftsriese – Bildungszwerg. Der Diskussionshintergrund zum Bildungsgesamtplan 1973. Analysen des OECD-Reports*. Reinbek: Rowohlt
- Maier, Harry 1994: *Bildungsökonomie. Die Interdependenz von Bildungs- und Beschäftigungssystem*. Stuttgart: Schäffer-Poeschel
- OECD (Hrsg.) 2001: *Lernen für das Leben. Erste Ergebnisse der internationalen Schulleistungsstudie PISA 2000*. Paris: OECD Publications
- O'Leary, John; Owen, Glen 2001: *UK pupils near top of the world league*. In: *Times Educational Supplement TES*, 04.12.2001. Online unter: www.tes.co.uk (Stand: 2/02)
- Pelgrum Willem J.; Plomp Tjeerd 1991: *The Use of Computers in Education Worldwide, Results from the IEA 'Computers in Education' Survey in 19 Education Systems*. Oxford: Pergamon
- Pelgrum Willem J.; Plomp, Tjeerd (Hrsg.) 1993: *The IEA Study of Computers in Education: Implementation of an Innovation in 21 Education Systems*. Oxford: Pergamon 1993
- Pelgrum Willem J.; Janssen Reinen, Ingeborg A.M.; Plomp, Tjeerd (Hrsg.) 1993: *Schools, Teachers, Students and Computers: a Cross-National Perspective*. Enschede: IEA
- Plomp, Tjeerd 2001: *The Potential and Challenges of International Comparative Studies of Educational Achievement*. In: Leimu, Kimmo; Linnakylä, Pirjo; Välijärvi, Jouni (Hrsg.): *Merging national and international interests in educational system evaluation*. Jyväskylä: University of Jyväskylä
- Purves, Alan C. (Hrsg.) 1992: *The IEA Study of Written Composition II: Education and Performance in Fourteen Countries*. Oxford, New York, Seoul, Tokyo: Pergamon
- Schultze, Walter; Riemenschneider, Lothar 1967: *Eine vergleichende Studie über die Ergebnisse des Mathematikunterrichts in zwölf Ländern*. Frankfurt a.M.
- Slater, John: *Shocking news – we are doing OK*. In: *Times Educational Supplement TES*, 07.12.2001. Online unter: www.tes.co.uk (Stand: 2/02)
- Torney-Purta, Judith u.a. (2001): *Citizenship and Education in Twenty-Eight Countries. Civic Knowledge and Engagement at Age Fourteen*. Amsterdam: IEA
- Vuorio-Lehti, Minna 1998: *The Finnish national matriculation examination: The gateway to higher education*. Turku. Online unter: www.leeds.ac.uk/educol/ (Stand: 2/02)

Isabell van Ackeren, geb. 1974; Studium der Biologie, Germanistik und Erziehungswissenschaft für die Sekundarstufen I und II; Mitarbeiterin in der Arbeitsgruppe Bildungsforschung/Bildungsplanung an der Universität Essen; Anschrift: Universität Essen, Fachbereich 2, AG Bildungsforschung/-planung, 45117 Essen; E-mail: i.vanackeren@uni-essen.de

Dem Einleitungs- und dem Schlussteil liegen Textentwürfe von Klaus Klemm zugrunde.