

Die Geschichte der „Tabelle 1“

Das Projekt „Schulleistung“ des Max-Planck-Institutes für Bildungsforschung (MPIB) war nach Meinung von Fachleuten die aufwendigste und gründlichste Leistungsstudie, die je in Deutschland durchgeführt worden ist. Sie hat etliche Millionen D-Mark gekostet.

Hintergrundinformationen: Das „Max-Planck Institut für Bildungsforschung (MPIB)“ wurde 1964 von Dr. jur. h.c. Hellmut Becker gegründet, um für Theorie und Praxis des Unterrichts eine empirisch abgesicherte Basis zu schaffen. Professor Peter M. Roeder war von 1973 bis 1995 Direktor am MPIB und dort seit 1982 zuständig für den „Forschungsbereich Schule und Unterricht“. Er war auch Mitglied in jenem Unterausschuss „Experimentalprogramm“ des Deutschen Bildungsrates, der 1968 unter dem Vorsitz von Dr. Becker die *„Empfehlung für die Einrichtung von Schulversuchen mit Gesamtschulen“* konzipiert hat. Professor Jürgen Baumert ist seit 1996, als Nachfolger von Roeder, Direktor am MPIB. Professor Olaf Köller war dort bis 2002 sein Mitarbeiter. Er ist jetzt Geschäftsführender Direktor des „Leibniz-Institutes für die Pädagogik der Naturwissenschaften und Mathematik (IPN)“ an der Christian-Albrecht-Universität in Kiel.

Im Rahmen des MPIB-Projektes „Schulleistung“ wurde zwischen 1968 und 1970 in den zehn Ländern der damaligen Bundesrepublik und in West-Berlin an etwa 450 Gymnasien der Lernfortschritt von etwa 14.000 Gymnasiasten untersucht. Mehrköpfige Test-Teams waren je drei Tage am Anfang und am Ende des 7. Jahrgangs damit beschäftigt, die Daten zum Leistungsstand für Deutsch, Englisch und Mathematik und weitere Auskünfte einzuholen.

Von den beteiligten Schülern gab es neben Informationen zu ihrem Freizeitverhalten auch Informationen zu ihrer kognitiven Grundfähigkeit (nach IST-Amthauer) und zu ihrem familiären Hintergrund. Der war ermittelt worden anhand des Bildungsstandes der Eltern- und der Großelterngeneration. Bei gleicher Gelegenheit wurden von etwa 1.130 Lehrern Auskünfte zu ihren Unterrichtsstrategien eingeholt. Die so gewonnene Datenbasis bestand aus *„verarbeitbaren Datensätzen für 12.594 Schüler aus 427 Klassen und für insgesamt 1.130 Deutsch- Englisch- und Mathematiklehrer“*. (Baumert u.a., 1986, S.644)

Wolfgang Edelstein vermerkt 1970 in seiner Beschreibung des Projektes: *„Die Kooperation aller Kultusministerien und des Deutschen Philologenverbandes verdient dankbare Erwähnung. Ohne sie wäre die Untersuchung kaum möglich gewesen.“* (1970, S.517). Offenbar ist von Seiten der Kultusministerien noch nie nach den Ergebnissen dieser „Kooperation“ gefragt worden.

Im Jahre 1965 hatte Wolfgang Edelstein, seinerzeit Assistent von Dr. Becker, das Projekt in der Odenwaldschule in kleiner Runde vorgestellt. Mit dabei waren Peter M. Roeder, Fritz Sang, Werner Stegelmann, Hartmut Zeiher und Jürgen Habermas als Mitglied des Institutsbeirates.

Gegenstand der Untersuchung waren die *„Determinanten von Leistung in der Schule“*. Es sollte herausgefunden werden, welche *„Faktoren des Systems“* Schulleistungen fördern und welche sie behindern (Edelstein, 1970, S.517 und 519). Weil in dieser Studie ausschließlich die Leistungsentwicklung von Gymnasiasten untersucht worden war, wird sie auch *„Gymnasiastenstudie“* genannt.

Genauere Auskünfte zum eigentlichen Gegenstand der Studie, zu den *„Determinanten von Leistung in der Schule“* gab es aus dem MPIB erstmals im Jahre 1986, und zwar in dem von Jürgen Baumert, Peter M. Roeder, Fritz Sang und Bernhard Schmitz veröffentlichten Aufsatz über *„Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen“*.

„*Leistungsentwicklung*“ ist hier ein Tauschwort für „Leistungsförderung“. Der Begriff „Leistungsausgleich“ meint die Verringerung des Abstandes zwischen dem oberen Leistungsfeld und dem unteren Leistungsfeld. Er wird in der Fachsprache „Streuung“, „Varianz“ oder „Divergenz“ genannt.

Die Vereinbarkeit von Leistungsförderung und Leistungsausgleich ist das zentrale Problem jeden Unterrichts. Es geht darum, „*Leistungsunterschiede zwischen Schülern ausgleichen zu wollen, ohne zugleich leistungsstärkere Schüler zu benachteiligen*“ (Baumert u.a., 1986, S.639). Und genau dies ist das Problem der Gesamtschulen und aller anderen Organisationsformen des Unterrichtes, in denen über den 4. Jahrgang hinaus in undifferenzierten Lerngruppen unterrichtet wird.

Die wichtigste Auskunft zur Vereinbarkeit von Leistungsförderung und Leistungsausgleich war die, dass es schon am Gymnasium für eine erfolgreiche Leistungsförderung deutliche Grenzen der Heterogenität gibt. Am effektivsten ist dort jener Unterrichtsstil, dessen Merkmale „*ein hohes Anspruchsniveau und ein zügiges Fortschreiten im Stoff*“ sind (Baumert u.a. 1986, S.655).

Wenn jedoch in einer Klasse die Unterschiede der Vorkenntnisse und der Begabungen allzu groß sind, dann zwingt das die Lehrer zu einer Verlangsamung des Unterrichtstempos sowie zu verstärktem Üben und Wiederholen.

„*Diese repetitive Unterrichtsführung nützt wider Erwarten Schülern mit ungünstigen Eingangsbedingungen nur wenig, während die Lernfortschritte der Schüler des oberen Leistungsdrittels merklich beeinträchtigt werden.*“ (Baumert u.a. 1986, S.655)

„*In nach Leistungsgruppen getrennten Analysen konnte gezeigt werden, dass bei streuungsverringendem Unterricht erhebliche Einbußen im Lernfortschritt des oberen Leistungsdrittels relativ schmalen Gewinnen im unteren Leistungsdritteln gegenüberstehen.*“ (Baumert u.a. 1986, S.654)

Fazit: Vergleiche „vergleichbarer“ Schüler haben also gezeigt: Jener Unterrichtsstil, dessen Merkmale „*ein hohes Anspruchsniveau und ein zügiges Fortschreiten im Stoff*“ sind, bringt den höchsten „Lernzuwachs“. Ein Unterricht aber, der auf Leistungsausgleich angelegt ist oder - wie an Gesamtschulen - auf Leistungsausgleich angelegt sein muss, geht stets zu Lasten der leistungsstärkeren Schüler. Er bringt keine so großen Vorteile für die leistungsschwächeren Schüler, dass deswegen eine Benachteiligung der leistungsstärkeren Schüler verantwortet werden könnte.

„*Selbst unter den günstigen Bedingungen des Gymnasiums ist es eine schwierige Aufgabe, eine befriedigende Balance zwischen optimaler Förderung und Leistungsausgleich zu finden. Sie wird nur von einer Minderheit der Lehrer wirklich bewältigt.*“ (Roeder/Sang 1991, S.164)

Weitere Informationen, zum Beispiel zum Thema „Binnendifferenzierung“, sind zu finden auf www.schulformdebatte.de unter „MPIB-Projekte“: MPIB-Projekt „Hauptschule/Gesamtschule“.

Mit ihrem Aufsatz „*Über die institutionelle Verarbeitung von Leistungsunterschieden*“ von 1991 liefern Roeder und Sang eine genauere Erklärung, warum es schon in Gymnasialklassen bei allzu großer Heterogenität zu einer Verlangsamung des Lerntempos kommt. Sie verweisen dazu auf eine schwedische Untersuchung von Ulf P. Lundgren zur Rolle der so genannten „*Steuerungsgruppe*“. Lundgren ist ein schwedischer Experte für die Rahmenbedingungen des Unterrichtens („*frame factors*“) und Professor an der Universität Uppsala.

Lehrer orientieren sich, so fand Lundgren heraus, in ihrem Unterrichtsverhalten „*in vieler Hinsicht an einer Gruppe von Schülern, die in der Intelligenzverteilung innerhalb der Klasse zwischen dem 10. und dem 25. Perzentil angesiedelt sind.*“ Die „Steuerungsgruppe“ rangiert also in der oberen Hälfte des untersten Viertels (Roeder/Sang, S.163).

„*Diese Beobachtung impliziert, dass die Lehrer ihre Aufmerksamkeit auf die Schüler konzentrieren, deren Leistungen zumeist deutlich unter dem Durchschnitt liegen. Ihnen ist es wichtiger, dass im Unterricht niemand auf der Strecke bleibt, als die ohnehin leistungsstarken Schüler ihren Möglichkeiten entsprechend zu fördern. Zu erwarten wäre also eine Angleichung der Schülerleistungen in diesen Klassen, wobei die Reduktion der Leistungsvarianz auf Kosten der leistungsstarken Schüler erreicht wird. Diese Erwartung wird durch eine ältere Untersuchung des MPIB bestätigt (Baumert, Roeder, Sang & Schmitz, 1986).*“ (Roeder/Sang 1991, S.163; es ist dies der bereits erwähnte Aufsatz über Leistungsförderung und Leistungsausgleich in Gymnasialklassen.)

Die „Tabelle 1“

Im Anschluss an diese Ausführungen referieren Roeder und Sang anhand einer Tabelle die Ergebnisse eines Vergleichs, den sie mit Daten des MPIB-Projektes „Schulleistung“ durchgeführt hatten. Die Tabelle trägt den Namen „Tabelle 1“.

Grundlage des Vergleichs war erstens der Leistungsstand, der am Anfang des 7. Jahrgangs bei jenen 1.049 Gymnasiasten vorgefunden worden war, die bis dahin sechsjährige Grundschulen besucht hatten. Das waren Klassen mit großen Leistungsunterschieden und hoher Streuung. Die andere Vergleichs-Grundlage war der Leistungsstand, der am Anfang des 7. Jahrgangs bei den mehr als 11.000 Gymnasiasten vorgefunden worden war, die in den übrigen Bundesländern nach vierjähriger Grundschule schon seit zwei Jahren Gymnasien besuchen konnten.

Ein Hinweis zum Verständnis der Zahlenwerte: „Die Testleistungen sind in Prozentpunktwerten angegeben: Unterschiede können also als Differenzen in Prozent des maximal erreichbaren Testwertes interpretiert werden.“ (Roeder/Sang 1991, S.166)

Tab. 1. Leistungsunterschiede und Unterschiede der Leistungsentwicklung zwischen Gymnasialschülern des 7. Schuljahres aus sechsjährigen Grundschulen und solchen aus vierjährigen Grundschulen plus zwei Jahre Gymnasium (N = 12.252 — insgesamt)

		Englisch	Mathematik	Deutsch	N
Anfang des 7. Schuljahres	sechsjährige \bar{x}	38.8	38.0	49.5	1.049
	Grundschule s	9.4	11.9	11.6	
	vierjährige \bar{x}	52.6	47.1	54.8	11.203
	Grundschule s	12.4	13.9	12.6	
Ende des 7. Schuljahres	sechsjährige \bar{x}	42.0	40.9	39.7	1.049
	Grundschule s	10.1	11.9	10.4	
	Residuen	-.80	.70	-.84	
	vierjährige \bar{x}	53.9	46.0	44.5	11.203
	Grundschule s	12.8	14.2	11.9	
	Residuen	.08	-.07	.08	

Roeder und Sang kommentieren die Tabelle folgendermaßen:

„Die Unterschiede zwischen den beiden Gruppen, die wir hier vergleichen - Gymnasiasten nach sechsjähriger Grundschule und solcher nach vierjähriger Grundschule plus zwei Jahre Gymnasium -, sind nicht nur höchst signifikant, was bei einer Stichprobe von mehr als 12.000 Schülern kaum überraschen kann, sondern in einer Größenordnung, die auch praktisch bedeutsam ist. Im Englischen beträgt der durchschnittliche Leistungsunterschied zu Beginn des Schuljahres etwas mehr als eine Standardabweichung, im Mathematikunterricht etwa eine dreiviertel Standardabweichungen, im Deutschen etwas weniger als eine halbe Standardabweichung.“ (Roeder/Sang 1991, S.166/167)

Auf Anfrage erhielten wir beim MPIB die Auskunft: In Mathematik und den Naturwissenschaften entspricht eine Drittel-Standardabweichung dem Lernfortschritt eines Schuljahres. Im Fach Englisch gilt das erst für eine halbe Standardabweichung.

Das bedeutet: In den 5. und 6. Jahrgängen der Gymnasien erreichten die Schüler in beiden Hauptfächern ein doppelt so hohes Lerntempo wie ihre Altersgenossen in den 5. und 6. Jahrgängen von sechsjährigen Grundschulen. Infolgedessen hatten die Gymnasiasten der Länder mit vierjähriger Grundschule am Anfang des 7. Jahrgangs in Mathematik und Englisch im Durchschnitt einen Wissensvorsprung von etwa zwei Schuljahren.

Mit anderen Worten: Die Schüler aus Berlin und Bremen hätten bei ihrem bisherigen Lerntempo erst nach zwei weiteren Schuljahren jenen Leistungsstand erreicht, den die Gymnasiasten aus Ländern mit vierjähriger Grundschule im Durchschnitt ihrer Gesamtheit am Anfang des 7. Jahrgangs bereits hatten. Nur in den homogeneren Klassen der 5. und 6. Jahrgänge der Gymnasien konnte jener Unterrichtstil zur vollen Wirkung kommen, dessen Merkmale „*ein hohes Anspruchsniveau und ein zügiges Fortschreiten im Stoff*“ sind.

Als Absicherung dieser Interpretation der Standardabweichungen nachfolgend ein Kommentar aus einer jüngeren Veröffentlichung des MPIB: „*Vergegenwärtigt man sich den Befund, dass eine Leistungsdifferenz von 0,3 bis 0,5 Standardabweichungen in der Mittelstufe mit dem Wissenszuwachs eines Schuljahres korrespondiert, so wird das Ausmaß und die praktische Bedeutung der Leistungsunterschiede erkennbar.*“ (Köller, Baumert u.a. 2004, S.691)

Die „Tabelle 1“ enthält auch präzise Angaben zu dem Leistungsfortschritt, der dann ab Anfang des 7. Jahrgangs bis zum Ende des 7. Jahrgangs erreicht wurde: „*In der letzten Zeile sind die Residuen des Leistungszuwachses unter Kontrolle der Ausgangsleistung, der Intelligenz und des Familienbackgrounds wiedergegeben. Sie sind ein möglichst direktes Maß für die Förderungsleistung des Unterrichts im 7. Schuljahr des Gymnasiums.*“ (Roeder/Sang 1991, S.166)

Mit diesem Rechenverfahren kann anhand von sogenannten „Regressions-Residuen“ die reine Effektivität des Unterrichtens beschrieben werden, unabhängig von den Vorkenntnissen der Schüler, ihrer kognitiven Grundfähigkeit und der zu erwartenden Unterstützungsleistung des Elternhauses.

„*Wie die Residualwerte für den Leistungszuwachs belegen, bleiben die Klassen nach sechsjähriger Grundschule im Englischen und im Deutschen etwas hinter dem durchschnittlichen Leistungsfortschritt zurück; in Mathematik liegen sie etwas darüber. Diese Unterschiede sind zwar statistisch, aber kaum noch praktisch signifikant. Beide Schülergruppen machen insgesamt durchaus vergleichbare Lernfortschritte, wodurch freilich auch die zu Beginn des Jahres bestehenden Leistungsunterschiede weitgehend erhalten bleiben.*“ (Roeder/Sang 1991, S.167)

Mit anderen Worten und in vereinfachender Darstellung: Das Lerntempo ist gleich, aber die Leistungsunterschiede erfahren keinen Ausgleich, sondern lediglich eine Parallel-Verschiebung. Wenn das so ist, dann können die Leistungsdefizite, die bei undifferenziertem Unterricht in den 5. und 6. Jahrgängen unvermeidlich entstehen, mit **PISA-Daten im 9. Jahrgang noch nachgewiesen** werden.

„*Die Aussagen der Lehrer über ihre Unterrichtsplanung für das 7. Schuljahr und über die Themen und Aufgaben, die sie späteren Schuljahren zuweisen, deuten darauf hin, dass der Leistungsausgleich zwischen beiden Gruppen langsamer vonstatten gehen wird, als man bei gleich befähigten Schülern vermuten würde.*“ (Roeder/Sang 1991, S.167/168)

Fazit: Schon am Gymnasium ist die Vereinbarkeit von Leistungsförderung und Leistungsausgleich ein großes Problem. Schon hier gibt es für einen effizienten Unterricht bei der Zusammensetzung der Klassen erkennbare Grenzen der Heterogenität. Um wie viel mehr dann in den 5. und 6. Jahrgängen von Schulen, an denen für diese Jahrgänge eine heterogene Zusammensetzung der Klassen zum Prinzip erhoben wurde. Weil dann der Unterricht stärker auf den Leistungsausgleich als auf die Leistungs-förderung angelegt sein muss, ist - zumindest in Deutschland - an Einheitschulen eine Benachteiligung der leistungsstärkeren Schüler unvermeidbar.

Dass auch die leistungsschwächeren Schüler in den leistungsgemischten Lerngruppen der 5. und 6. Jahrgänge von integrativen Schulen durch den Bezugsgruppen-Effekt hohen psychischen Belastungen ausgesetzt sind, die ihnen an Hauptschulen erspart bleiben, war schon 1984 einer der am besten bezeugten Befunde der Bildungsforschung. (s. www.schulformdebatte.de, unter „Hauptschule“)

Bei einer öffentlichkeitswirksamen Bekanntgabe dieser Forschungsergebnisse hätten ab 1991 in Deutschland keine Integrierten Gesamtschulen und keine anderen Schulen mehr gegründet werden dürfen, an denen über den 4. Jahrgang hinaus in undifferenzierten Lerngruppen unterrichtet wird.

Dann wäre schon das Jahr 1991 „Das Jahr der großen pädagogischen Wende“ gewesen.

Der erste Satz der „Schlussbemerkung“ zu dem Aufsatz von Roeder und Sang aus dem Jahre 1991 lautete jedoch: „*Wir wollen diesen Aufsatz bewusst nicht mit einem Plädoyer für die vierjährige Grundschule beschließen*“.(S.168) Er ist in diesem Kontext ebenso unerwartet wie unverständlich. Denn er steht in offenem Widerspruch zu den vorgelegten Forschungsergebnissen und zu den weiteren Mitteilungen der Schlussbemerkung. Weil dieser Satz seitdem eine wirklichkeitsnahe Mitteilung von Ergebnissen der Bildungsforschung behindert hat, haben wir die gesamte Schlussbemerkung als Kopie in den Text aufgenommen.

Schlußbemerkung

Wir wollen diesen Aufsatz bewußt nicht mit einem Plädoyer für die vierjährige Grundschule beschließen, schon weil sich auch andere schulpolitische Folgerungen aus unseren Ergebnissen ziehen lassen. Und während der systematische Ertrag der Analyse relativ unabhängig vom Zeitpunkt der Erhebung sein dürfte, sollten sich schulpolitische Schlußfolgerungen zweifellos auf neuere Erhebungsdaten stützen.

Systematisch kann der Übergang aufs Gymnasium nach vier Schuljahren im Vergleich zur sechsjährigen Grundschule als ein "advanced placement" bewertet werden, das an vielen amerikanischen High Schools praktiziert wird. Dabei werden Schüler mit besonders hohen Leistungen in Lerngruppen aufgenommen, in denen wesentlich höhere Lernanforderungen gestellt werden. Während Metaanalysen zur äußeren Leistungsdifferenzierung insgesamt zu eher diffusen Ergebnissen führen, zeigt die empirische Evaluation von "advanced placement" wie die hier vorgestellte Untersuchung deutlich überdurchschnittliche Leistungszuwächse bei den betroffenen Schülern (Fraser, Walberg, Welch & Hattie, 1987; Kerckhoff, 1986; Natriello, Pallas & Alexander, 1989).

Kommentar: Der Begriff „advanced placement“ dient an amerikanischen Schulen als Bezeichnung für eine Rangstellung. Es sind das Schüler, denen wegen ihrer überdurchschnittlichen Kenntnisse und Fähigkeiten anspruchsvollere Aufgaben zugewiesen werden. Hier werden mit dieser Bezeichnung Schüler bewertet, die nach vier Jahren Grundschule schon seit zwei Jahren Gymnasien besuchen konnten.

Der erste Satz der Schlussbemerkung ist ein klares Bekenntnis zu der bis dahin geübten „zurückhaltenden“ Informations-Politik des MPIB, aber in dieser Deutlichkeit neu - und daher nur schwer zu widerrufen. Damit war eine Spur gelegt worden, die nicht mehr verlassen werden konnte, ohne das Ansehen Roeders und die Reputation des Institutes zu beschädigen.

Im Rückblick und ins Präteritum gesetzt, wirkt der Satz fast wie ein Geständnis: *Wir wollten diesen Aufsatz bewusst nicht mit einem Plädoyer für die vierjährige Grundschule beschließen.*

Die Frage, ob die Widersprüchlichkeiten der „Schlussbemerkung“ nun ein Zeichen von Unentschlossenheit oder das Produkt anderer Motivationen sind, ist inzwischen relativ unwichtig. Roeder gehörte 1968 immerhin zu den Initiatoren der „Schulversuche mit Gesamtschulen“.

Viel wichtiger ist die Tatsache, dass dieser Aufsatz - nach solchen Vorarbeiten und angesichts der hier gewonnenen Erkenntnisse - eigentlich nur mit einem entschiedenen „Plädoyer für die vierjährige Grundschule“ hätte enden dürfen.

„Nicht geschehene Taten lösen einen katastrophalen Mangel an Folgen aus.“

(Stanislaw Jerzy Lec, polnischer Autor, 1909-1966)

Dass sich schulpolitische Schlussfolgerungen „zweifellos auf neuere Erhebungsdaten stützen“ sollten, ist ein klarer Hinweis auf das MPIB-Projekt „Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU)“. Es wurde 1991 unter der Leitung von Roeder und Baumert gestartet, ebenfalls im 7. Jahrgang, als „Teilwiederholung“ des MPIB-Projektes „Schulleistung“ (s. BIJU 1996, S.8, erste Zeile), allerdings nur in vier Bundesländern (darunter NRW und Berlin), aber in Schulen aller Schulformen, auch in Gesamtschulen.

Erkenntnisse über **Leistungsentwicklung und Ausgleich von Leistungsunterschieden an Gesamtschulen** hätten mit den BIJU-Daten aus NRW und Berlin schon 1998 vorgelegt werden können. In jenem „für ein breiteres Publikum gedachten Bericht“ über die BIJU-Befunde, der von Baumert und Köller im Juni zum Jahresende 1998 angekündigt worden war, hätte das schon vor mehr als zehn Jahren geschehen können (vgl. „Pädagogik“ 6/1998, S.13, oben links).

Die „Tabelle XII.1“

In der 1997 von Franz E. Weinert und Andreas Helmke herausgegebenen Aufsatzsammlung „Entwicklung in Grundschulalter“ hatte Roeder die Gelegenheit einer Korrektur. Er veröffentlichte dort die „Tabelle 1“ noch einmal, jetzt als Tabelle XII.1, mit der gleichen Überschrift („Leistungsunterschiede und“) und mit dem gleichen Zahlenspiegel (Roeder 1997, S.408).

Tabelle XII.1: Leistungsunterschiede und Unterschiede der Leistungsentwicklung zwischen Gymnasialschülern des 7. Schuljahres aus sechsjährigen Grundschulen und solchen aus vierjährigen Grundschulen plus zwei Jahre Gymnasium (N = 12.252 – insges.)

			ENGLISCH	MATHEMATIK	DEUTSCH	N
Anfang des 7. Schuljahres	sechsjährige Grundschule	x	38.8	38.0	49.5	1.049
		s	9.4	11.9	11.6	
	vierjährige Grundschule	x	52.6	47.1	54.8	11.203
		s	12.4	13.9	12.6	
Ende des 7. Schuljahres	sechsjährige Grundschule	x	42.0	40.9	39.7	1.049
		s	10.1	11.9	10.4	
		Residuen	-.80	.70	-.84	
	vierjährige Grundschule	x	53.9	46.0	44.5	11.203
		s	12.8	14.2	11.9	
Residuen	.08	-.07	.08			

Legende: x = Mittelwert; s = Standardabweichung

Zu Beginn der Auswertung (Roeder 1997, S.407) hatte es dort noch geheißen: „Ein wesentliches Ergebnis dieser Sekundäranalyse war, dass Lehrer, die ihre Schüler nach sechsjähriger Grundschule aufnahmen, im Vergleich zu den Fachlehrern an grundständigen Gymnasien deutlich niedrigere Erwartungen hinsichtlich der Vorkenntnisse ihrer Schüler hatten und ihren Unterricht im 7. Jahrgang entsprechend planten. Wie Tabelle XII.1 zeigt, bestätigen die Testleistungen der Schüler zu Beginn des 7. Schuljahres die Erwartungen der Lehrer.“ (Unterstreichungen wurden nachträglich eingefügt. - Eine Primär-Analyse des Projektes „Schulleistung“ hat es übrigens nie gegeben.)

„Die Leistungen nach sechsjähriger Grundschule liegen erheblich unter denen von Schülern, die den Wechsel aufs Gymnasium bereits nach der vierten Grundschulklasse vollzogen haben. Für Englisch und Mathematik beträgt der Abstand etwa eine Standardabweichung, für Deutsch eine halbe. Am Ende des 7. Schuljahres haben sich die Leistungen beider Schülergruppen zwar ein Stück weit angeglichen, aber die verbleibenden Unterschiede sind - vor allem im Englischen - immer noch bedeutsam.“ (Roeder 1997, S.407;408; - 1991 war es erheblich weniger als „ein Stück weit“).

Zunächst werden die Leistungsunterschiede noch sehr deutlich als eine Folge der verspäteten Differenzierung beschrieben. Aber dann werden ihre Ursachen wieder „verschleiert“: „Sie lassen sich, was ich hier nicht ausführen will, wohl im Wesentlichen auf die curricularen Entscheidungen der beiden Lehrergruppen zurückführen.“ (Roeder 1997, S.408)

So entsteht der Eindruck es handele sich hier um einen „Anti-Pygmalion-Effekt“: Die Schüler leisteten nicht mehr, weil ihre Lehrer im 7. Jahrgang nicht mehr erwarteten. Gelegentlich hat man den Eindruck, als solle bzw. müsse mit derartigen Formulierungen *eine Zensur des Zeitgeistes* unterlaufen - oder überlistet werden.

Die „Tabelle 23.2“

Im Jahre 2002, also 32 Jahre nach Abschluss der Datenerhebung, hatten die Professoren Baumert und Köller als erste den Mut, die Daten der „Tabelle 1“ zu einem entschiedenen „*Plädoyer für die vierjährige Grundschule*“ zu nutzen. Das geschah im Lehrbuch der „Entwicklungspsychologie“ von Oerter/ Montada (2002, S.770).

Sie beziehen sich dabei auf die von Roeder 1997 veröffentlichte „Tabelle XII.1“. Doch wird von ihnen in ihrer „Tabelle 23.2“ lediglich der obere Teil der „Tabelle XII.1“ übernommen, ohne die Daten zu den am Ende des 7. Jahrgangs vorgefundenen Leistungen und ohne die „Residuen“ mit ihren Informationen zu der im 7. Jahrgang weiterwirkenden Benachteiligung der Schüler von sechsjährigen Grundschulen.

Der Abschnitt steht unter der Überschrift: „*Frühe Differenzierung fördert leistungsstarke Schüler*“. Zunächst teilen Köller und Baumert mit, dass in anderen Ländern eine fähigkeitsbezogene Gruppierung der Schüler nach Art des gegliederten Schulwesens keine Vorteile hat. Für Deutschland gelte dies jedoch nicht.

Sie verweisen dann auf jene von Roeder und Sang veröffentlichte Auswertung der Daten des MPIB-Projektes „Schulleistung“ und beschreiben die Problemstellung des von Roeder und Sang veröffentlichten Vergleichs: Es ging um die zentrale Frage aller Schulform-Debatten, um die Frage, „*ob eine spätere Differenzierung (nach der 6. Jahrgangsstufe) bei leistungsstarken Schülern im Vergleich zur Differenzierung nach vier Schuljahren ungünstige Effekte auf die Schulleistungen hat. Dazu wurden die Leistungen von Gymnasiasten in der 7. Jahrgangsstufe in Deutsch, Mathematik und Englisch aus einem Bundesland mit sechsjähriger Grundschule mit denen der übrigen Bundesländer*

mit vierjähriger Grundschule verglichen. Die Tabelle 23.2 zeigt die Resultate der Studie.“

Tabelle 23.2. Schulleistungen in der 7. Jahrgangsstufe von Gymnasiasten aus vier und sechsjährigen Grundschulen (aus Roeder, 1997, S. 408)

		Englisch	Mathematik	Deutsch	N
sechsjährige Grundschule	\bar{x}	38.8	38.0	49.5	1049
	<i>s</i>	9.4	11.9	11.6	
vierjährige Grundschule	\bar{x}	52.6	47.1	54.8	11.203
	<i>s</i>	12.4	13.9	12.6	

Anmerkungen. \bar{x} : Mittelwert; *s* = Standardabweichung

Diese Tabelle 23.2 wird von Köller und Baumert folgendermaßen kommentiert: „*Unübersehbar zeigen sich deutliche Leistungsvorteile der Gymnasiasten aus den Bundesländern mit vierjähriger Grundschule. Die frühere Differenzierung scheint also leistungsstärkere Schüler in allen drei Fächern besser zu fördern.*“ (S.770)

Der nachfolgende Exkurs schließt mit dem Satz: „Zusammenfassend zeigt sich zumindest für das deutsche Schulsystem, dass bezogen auf die Fachleistungsentwicklung leistungsstarke Schüler von der Differenzierung im Sekundarbereich profitieren.“ (S.771) Jenes Bundesland mit sechsjähriger Grundschule war Berlin.

Dieses entschiedene Plädoyer für die vierjährige Grundschule erfolgte allerdings an abgelegener Stelle, in der Fachliteratur, und nicht mit jener Öffentlichkeitswirkung, die seiner Bedeutung und dem Aufwand seiner Erstellung entsprochen hätte. Es ist selbst in der Fachwelt kaum bekannt.

Nur so war es möglich, dass im „KMK-Bildungsbericht 2003“ just diese Hinweise auf die Vorteile einer frühen Differenzierung, die von Köller und Baumert im Oerter/Montada 2002 veröffentlicht worden waren, innerhalb der sonst wortgetreuen Wiedergabe des Kontextes unbemerkt ausgeblendet werden konnten.

Der „KMK-Bildungsbericht 2003“ wurde herausgegeben vom „Deutschen Institut für Internationale Pädagogische Forschung (DIPF)“. Wir haben mit einem Offenen Brief an Herrn Professor Dr. Eckhard Klieme, den Direktor des DIPF, vom 8. Januar 2008 auf diese „Ausblendung“ hingewiesen (vgl. www.schulformdebatte.de unter „Archiv“: „Sollte das gegliederte Schulwesen zum Opfer einer Intrige werden?“). Der Brief blieb ohne Reaktionen.

Die schulart-unabhängig Orientierungsstufe in Mecklenburg-Vorpommern, die sechsjährige Primarstufe in Hamburg und die fünfjährige Grundschule im Saarland sind durch die „Ausblendung“ erst möglich geworden. Auch der Streit um die ELEMENT-Studie wäre anders verlaufen.

Im Oerter/Montada 2008 (S.750) wurde dann, sechs Jahre später, der Versuch unternommen, die Bedeutung dieser wichtigen Erkenntnisse des MPIB zum niedrigeren Fördereffekt der 5. und 6. Jahrgänge von Berliner Grundschulen nachträglich durch eine Fußnote zu relativieren:

„Problematisch für die Interpretation der Befunde ist, dass hier lediglich ein Bundesland gegen alle übrigen kontrastiert wird und sonstige Besonderheiten der jeweiligen Schulstrukturen (beispielsweise unterschiedliche Expansionsraten des Gymnasiums) mit der Zahl der Grundschuljahre konfundiert sein können.“ (Oerter/Montada 2008, S.750)

Dabei hätten die Erkenntnisse zu den Vorteilen einer frühen Differenzierung an Hand der aus dem Projekt „Schulleistung“ vorliegenden Daten zum familiären Hintergrund der Schüler und zu ihren kognitiven Grundfähigkeiten ohne größeren Aufwand entsprechend präzisiert werden können.

Weil aus dem MPIB-Projekt *„Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU)“* für jeweils etwa 3.000 Schüler der 7. Jahrgänge des Jahres 1991 aus Berlin und NRW ebenfalls solche Daten vorliegen, hätten die MPIB-Befunde zu den „unübersehbaren“ Vorteilen einer frühen Differenzierung, wie Roeder das empfohlen hatte, schon 1996 mit „BIJU“-Daten präzisiert und aktualisiert werden können. Das ist, wie wir wissen und nachweisen können, auch geschehen. Aber die Ergebnisse sind am MPIB zurückgehalten worden.

Die Geschichte der „Tabelle 1“ ist also noch nicht zu Ende!

Wie von Professor Baumert und seinen Mitarbeitern mehrfach erläutert und am Beispiel der Laborschule Bielefeld eindrucksvoll demonstriert wurde, sind auch mit den PISA Daten der etwa 30.000 deutschen Neuntklässler Schulformvergleiche möglich (z.B. PISA 2000/01, S.128 und PISA 2000/06, S.116-123). Denn auch hier gibt es neben den Daten zu den familiären Verhältnissen der Schüler wiederum Daten zu ihrer kognitiven Grundfähigkeit. Aktuellere Daten als die aus PISA-E 2006 stehen zur Zeit für bundesweite Schulformvergleiche nicht zur Verfügung. Derartige Schulformvergleiche sollten daher jetzt von Behörden, Parteien und Verbänden mit Nachdruck eingefordert werden, im Interesse der Leistungsfähigkeit und der Stabilität unserer Gesellschaft.

Weitere Informationen auf: