Arbeitskreis Gesamtschule e.V. **Am Lohtor 11**
Ulrich Sprenger Vorsitzender **45657 Recklinghausen**

betrifft: Bildung in Deutschland

Offener Brief
des Arbeitskreises Gesamtschule e.V.

An das Deutsche Institut für Internationale Pädagogische Forschung
An die Kultusminister der Länder der Bundesrepublik Deutschland
An die Bundesministerin für Bildung und Forschung
An das Max-Planck-Institut für Bildungsforschung, zur Kenntnisnahme

12. Oktober 2006

Sehr geehrte Damen und Herren,

anbei senden wir Ihnen unter Bezug auf die im Jahre 2003 und im Jahre 2006 erschienenen Bildungsberichte eine Studie mit Hinweisen auf Problemfelder der Bildungsforschung, für die zur Zeit genauere Informationen dringend benötigt werden.

Wir wollen mit dieser Studie gleichzeitig aufmerksam machen auf Sachverhalte, die nun schon seit etlichen Jahren eine Bestandsaufnahme der bestehenden Verhältnisse spürbar behindern, zum Schaden des deutschen Schulwesens.

Diese Behinderung einer Bestandsaufnahme geschieht zum Nachteil der Schülerschaft ebenso wie zum Nachteil der sie unterrichtenden Lehrerinnen und Lehrer und zum Nachteile der Bundesrepublik Deutschland.

Im Interesse der Allgemeinheit wenden wir uns an Sie mit der Bitte, dafür Sorge zu tragen, dass bei der Erarbeitung der nächsten Bildungsberichte und in den kommenden Schulform-Debatten die von uns zusammengestellten Hinweise Beachtung finden.

Bezüglich einer ausführlicheren Darstellung der hier beschriebenen Sachverhalte verweisen wir auf unsere in Kürze erscheinende Dokumentation *„Der unkontrollierte Verfall des deutschen Schulwesens“*.

Mit freundlichen Grüßen

Sprenger

Vorsitzender

Anlage

Professor Dr. Peter M. Roeders Gutachten für den Landtag von Sachsen-Anhalt von 1995

- Die Bildungsberichte von 2003 und 2006 -

Qualitätsentwicklung ohne genauere Qualitätskontrollen?

Hinweise auf dringend benötigte Informationen aus der Bildungsforschung

Vorbemerkungen

Der Beschluss der Kultusministerkonferenz und des Bundesministeriums für Bildung und Forschung, ab 2006 alle zwei Jahre von einem unabhängigen Konsortium Berichte zur *„Bildung in Deutschland“* veröffentlichen zu lassen, ist für die Zukunft von Theorie und Praxis des Unterrichtens von sehr hoher Bedeutung.

Diese Bildungsberichte sind als *„indikatorengestützte Berichte“* konzipiert. Das heißt, sie basieren auf überprüfbaren Daten. *„Sie müssen empirisch belastbare Informationen über einen relevanten Ausschnitt des Bildungswesens enthalten, sich auf fortschreibbare Datensätze stützen und bundesweite, soweit möglich auch international und länderspezifisch vergleichende Aussagen zulassen.“*

„Will Bildungspolitik in Deutschland mit den Mitteln einer indikatorengestützten Berichterstattung das Wissen über die Leistungsfähigkeit des Bildungswesens erhöhen und zugleich gezielt zur Beseitigung von dabei sichtbar werdenden Defiziten beitragen, so bedarf es in den nächsten Jahren erheblicher Anstrengung. Notwendig ist vor allem die Verbesserung der Datenbasis, die schrittweise dazu führt, dass weiter gehende Einsichten und Erkenntnisse für eine umfassende und zuverlässige Berichterstattung verfügbar sind.“ (Bildungsbericht 2006, S.2/3)

In den Bildungsberichten, die unter der Federführung des Deutschen Institutes für Internationale Pädagogische Forschung (DIPF) erstellt und in den Jahren 2003 bzw. 2006 veröffentlicht worden sind, ist bereits eine Fülle bildungs-relevanter Informationen veröffentlicht worden.

Gleichwohl werden zuverlässige, datengestützte und, wenn möglich, mit allgemeinverständlichen Zusammenfassungen versehene Auskünfte zu den nachfolgend beschriebenen Fragenkomplexen des „Forschungsbereiches Schule und Unterricht“ dringend benötigt.

Der Arbeitskreis Gesamtschule e.V. darf von sich behaupten, in den zwölf Jahren seines Bestehens zu der ohnehin vorhandenen praktischen Kompetenz auch eine relativ hohe theoretische Kompetenz erworben zu haben.

Daher erlauben wir uns, in den nachfolgenden Seiten auf Problemfelder hinzuweisen, für die unseres Wissens gründlichere, datengestützte Erkenntnisse aus der Bildungsforschung *„zur Beseitigung von Defiziten“* noch nicht vorliegen oder noch nicht veröffentlicht wurden oder noch nicht mit der erforderlichen breitenwirksamen Allgemeinverständlichkeit bekannt gemacht worden sind.

Die bestehenden Verhältnisse können im Sinne einer Qualitätsentwicklung erst dann optimiert werden, wenn zuvor für die einzelnen Bundesländer der Versuch unternommen wurde, mit Qualitätskontrollen und *„empirisch belastbaren“* Berichten auch eine verlässliche Darstellung des zu optimierenden *„Zustandes“*, also eine *„Erfassung“* der Stärken und Schwächen des Bestehenden vorzulegen, und wenn diese Darstellung der bestehenden Verhältnisse mit der erforderlichen breitenwirksamen Allgemeinverständlichkeit veröffentlicht worden ist.

Hinweise
auf dringend benötigte Informationen aus der Bildungsforschung

1. Es fehlen zur Zeit genauere Angaben zu der Benachteiligung, die ohnehin schon benachteiligten lernschwächeren Kindern in den undifferenzierten Klassen von weiterführenden Schulen widerfährt.

2. Es fehlen zur Zeit genauere Angaben zu den Vorteilen einer frühen, im 5. Jahrgang einsetzenden Differenzierung in getrennten Schulformen und zum niedrigen Fördereffekt des Unterrichts in den undifferenzierten Klassen der 5. und 6. Jahrgänge von sechsjährigen Grundschulen.

3. Es fehlt zur Zeit eine aktuelle, datengestützte Diskussion über die bereits 1984 von Helmut Fend beschriebenen Vorteile einer Dreier-Differenzierung - gegenüber der zur Zeit in manchen Bundesländern diskutierten Zweier-Differenzierung.

4. Es fehlen zur Zeit Berichte über die Ergebnisse der Vergleiche von Schulformen und Schulsystemen, die nach den *„Grundregeln der Schuleffektivitäts-Forschung“* als *„Vergleiche vergleichbarer Schüler“* auf drei unterschiedlichen, schulformbezogenen Vergleichs-Ebenen durchgeführt worden sind.

5. Es fehlen zur Zeit für die einzelnen Bundesländer Angaben zum unterschiedlichen Begabungs-Potential ihrer Schülerschaft.

6. Es fehlen genauere Informationen zur katastrophalen Situation der NRW-Gesamtschul-Oberstufen.

7. Es fehlt eine öffentliche Korrektur der im Teil C des KMK-Bildungsbericht 2003 enthaltenen, folgenreichen Unvollständigkeiten. Sie betreffen die Vorteile einer mit dem 5. Jahrgang einsetzenden frühen Differenzierung nach Fähigkeiten.

Kommentar

Bei den hier angesprochenen Bildungsberichten handelt es sich um den im Jahre 2003 erschienenen, im Auftrag der Ständigen Konferenz der Kultusminister erstellten *„Bildungsbericht für Deutschland“* und um den im Jahre 2006 erschienenen Bericht *„Bildung in Deutschland“*. Dieser Bildungsbericht aus dem Jahre 2006 ist im Auftrag der Ständigen Konferenz der Kultusminister und des Bundesministeriums für Bildung und Forschung erstellt worden, und zwar vom *„Konsortium Bildungsberichterstattung“*.

Die Federführung für beide Bildungsberichte lag beim *„Deutschen Institut für Internationale Pädagogische Forschung* (DIPF)*“*. Die Veröffentlichung weiterer Bildungsberichte - im Abstand von zwei Jahren - gehört zum Programm des Auftrags.

Der vollständige Titel des Bildungsberichtes 2003 lautet: *„Bildungsbericht für Deutschland - Erste Befunde“*. Die Autoren dieses Bildungsberichtes sind: Hermann Avenarius (Sprecher), Hartmut Ditton, Hans Döbert, Klaus Klemm, Eckhard Klieme, Matthias Rürup (Koordination), Heinz-Elmar Tenorth, Horst Weishaupt, Manfred Weiß, gemeinsam mit Hans-Peter Füssel, Olaf Köller und Rainer H. Lehmann. (Im nachfolgenden Text wird der Bildungsbericht 2003, um ihn vom MPIB-Bildungsbericht 2003 zu unterscheiden, unter der Bezeichnung „KMK-Bildungsbericht 2003“ geführt.)

„Die Kultusministerkonferenz hat im Februar 2002 beschlossen, die Öffentlichkeit systematisch, umfassend und kontinuierlich über wichtige Daten und Entwicklungen im Bildungswesen der Bundesrepublik Deutschland zu informieren.“ (Vorwort des Bildungsberichtes 2003, S.1)

Der vollständige Titel des Bildungsberichtes 2006 lautet: *„Bildung in Deutschland - Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration“*. Der Bildungsbericht 2006 wird folgendermaßen charakterisiert:

„Der Bericht ist eine datengestützte Analyse von Bildung in Deutschland; er verzichtet auf Wertungen und Empfehlungen. Seine Besonderheit liegt darin, dass er eine problemorientierte Darstellung auf der Grundlage von Indikatoren *ist. Indikatoren sind, wie internationale Erfahrungen zeigen, die geeignetste Form der Erfassung des Zustands und der Entwicklung des Bildungswesens.“* (Bildungsbericht 2006, S.2; nachträglich unterstrichen)

„Das Konsortium hat den Bericht unter Wahrung seiner wissenschaftlichen Unabhängigkeit in Abstimmung mit einer Steuerungsgruppe erarbeitet, die für die Auftraggeber - die Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland und das Bundesministerium für Bildung und Forschung - die Arbeit des Konsortiums begleitete. Das Projekt wurde von einem wissenschaftlichen Beirat unter dem Vorsitz von Prof. Dr. Jürgen Baumert unterstützt. Die Zusammenarbeit mit beiden Gremien hat sich für das Konsortium als sehr fruchtbar erwiesen.“ (Vorwort, S.V)

Dem *„Konsortium Bildungsberichterstattung“* gehören an: (federführend) das Deutsche Institut für Internationale Pädagogische Forschung (DIPF), das Deutsche Jugendinstitut (DJI), die Hochschul-Informations-System GmbH (HIS), das Soziologische Forschungsinstitut an der Universität Göttingen (SOFI), das Statistische Bundesamt (StBA) und die Statistischen Ämter der Länder (StLÄ).

Die *„Verantwortlichen Autoren“* des Bildungsberichtes 2006 sind: Hermann Avenarius (DIPF), Martin Baethge (SOFI), Hans Döbert (DIPF), Heinz-Werner Hetmeier (StBA), Eckhard Klieme (DIPF), Gisela Meister-Scheufelen (StLÄ), Thomas Rauschenbach (DJI), Andrä Wolter (HIS).

Dringend benötigte Informationen aus der Bildungsforschung

1. Es fehlen zur Zeit genauere Angaben zu der Benachteiligung, die ohnehin schon benachteiligten Kindern in den undifferenzierten Klassen von weiterführenden Schulen widerfährt.

Spätestens ab 1984 gehört es zu den am besten bezeugten Befunden der Bildungsforschung, dass leistungsschwächere Schülerinnen und Schüler in den undifferenzierten Klassen von weiterführenden Schulen durch den Bezugsgruppen-Effekt und die unvermeidbaren ständigen Aufwärts-Vergleiche permanenten Unzulänglichkeits- und Beschämungserfahrungen ausgesetzt sind, die ihren Altersgenossen an Hauptschulen erspart bleiben. *„Die Hauptschule hat eine selbstwertschützende Funktion.“* (TIMSS II 1997, S.175)

In diesem Zusammenhang verweisen wir auf unseren Aufsatz über *„Die fortschreitende Benachteiligung der ohnehin schon Benachteiligten“*. - Die Komplexität der Problematik ist bereits 1977 von Helmut Fend und Werner Specht folgendermaßen beschrieben worden:

„Während an integrierten Gesamtschulen für den Schüler als Bezugsgruppe der eigenen Selbsteinschätzung und Selbstevaluation die gesamte Altersgruppe organisatorisch vorgegeben ist, dürften soziale Vergleichsprozesse im gegliederten System eher schulformintern erfolgen.

Die Gruppe der leistungsschwächeren Schüler gerät an Gesamtschulen so in eine psychisch schwierige Lage, weil ihnen ihre faktische, ungünstige Stellung in der schulischen Statushierarchie tagtäglich von neuem vor Augen geführt wird, während im Vergleich dazu Hauptschüler innerhalb des traditionellen Schulsystems die Möglichkeit haben, im Verhältnis zu ihren Mitschülern relativ erfolgreich zu sein. Diese Möglichkeit haben zwar auch Gesamtschüler innerhalb ihrer Kurse. Die Erfahrung, die sie in den leistungshomogenen Gruppen machen, dürften bewusstseinsmäßig aber doch stets von jenen überlagert werden, denen sie in der integrierten Kerngruppe begegnen.

An dieser Stelle müssen wir uns fragen, ob wir nicht bereits aufgrund der bisherigen Ergebnisse die Hoffnung als gescheitert betrachten müssen, die sich an ein integriertes Schulsystem als Lösungsmöglichkeit für die brennenden Probleme der Hauptschule geknüpft haben. Angesichts des Anspruchs der Humanisierung der Schule, mit dem die Gesamtschulbewegung stets aufgetreten ist, muss die Tatsache besonders enttäuschen, dass offenbar gerade diejenige Schülergruppe am wenigsten von der neuen Schulstruktur profitiert, die dieser Humanisierung am meisten bedarf und die eigentlich auch als Hauptzielgruppe der organisatorischen und curricularen Innovationen gedacht war, nämlich jene der schwächeren und sozial benachteiligten Schüler, denen man im Rahmen einer integrierten Organisationsform besondere Fördermaßnahmen angedeihen lassen wollte.

Beunruhigend erscheint vor allem, dass die relativ ungünstige Lage der schwächeren Schüler an Gesamtschulen, wie sich in den o. a. Daten zeigt, offenbar keine historische Zufälligkeit darstellt, sondern im wesentlichen als Folge des Strukturmerkmals der Integration gesehen werden muss, was die Erfolgsaussichten kompensierender Maßnahmen verständlicherweise verringert.“

(Aufsatzsammlung „Hauptschule“, 1977 herausgegeben von der Redaktion „betrifft: Erziehung“, S.40f; die Unterstreichung im letzten Abschnitt ist nachträglich eingefügt worden.)

Fazit: Es fehlen zur Zeit datengestützte Darstellungen zu der Frage, wie sich das Selbstwertgefühl und die Leistungen von leistungsschwächeren Schülern in leistungsgemischten Lerngruppen entwickeln. Die dazu erforderlichen Daten aus dem MPIB-Projekt *„Bildungsverläufe und psychosoziale Entwicklung im Jugendalter* (BIJU)“ (1991-2001) liegen bereits vor (vgl. Olaf Köller, 2. BIJU-Bericht, 1996, S.23, *Abbildung 11: „Entwicklung des Selbstwertgefühls“*).

Eine datengestützte Darstellung aus dem Projekt BIJU zu der Frage, wie sich das Selbstwertgefühl und die Leistungen von leistungsstarken Gymnasiasten in sehr leistungsstarken Gymnasialklassen entwickeln, gibt es schon (Köller/Baumert 2001).

<u>2.</u> Es fehlen zur Zeit genauere Angaben zu den Vorteilen einer frühen, im 5. Jahrgang einsetzenden Differenzierung in getrennten Schulformen und zum niedrigen Fördereffekt des Unterrichts in den undifferenzierten Klassen der 5. und 6. Jahrgänge von sechsjährigen Grundschulen.

Solche Informationen sind dringend erforderlich. In Mecklenburg-Vorpommern wurde 2005 zum Beispiel unter dem Eindruck der vorherrschenden PISA-Interpretationen und in dem Glauben, auf diese Weise das dortige Schulsystem optimieren zu können, die schulart-unabhängige Orientierungsstufe eingeführt. Das hätte verhindert werden müssen - und es hätte verhindert werden können.

Zwischenbemerkung: Die 5. und 6. Jahrgänge sind bisher in allen, schulformübergreifenden deutschen Leistungsstudien ausgespart geblieben, obgleich hier die Weichenstellungen jeder Schullaufbahn erfolgen.

Professor Dr. Peter M. Roeder, von 1973-1995 Direktor am MPIB, liefert dazu in seinem Aufsatz über *„Entwicklung vor, während und nach der Grundschulzeit“* (in: F. E. Weinert / A. Helmke: *Entwicklung im Grundschulalter*, S.407) folgenden Kommentar: *„Hinsichtlich der Schulleistungs-Effekte einer Verlängerung der Grundschuldauer liegt meines Wissens nur eine repräsentative Studie vor, nämlich eine Sekundäranalyse der um 1970 durch das Max-Planck-Institut für Bildungsforschung erhobenen Daten zur Leistungsentwicklung von Gymnasiasten des 7. Schuljahres (Roeder & Sang 1991).“*

Die Daten dieser *„Sekundäranalyse“* stammen aus dem MPIB-Projekt *„Schulleistung“* (1968-1970). Eine Primär-Analyse dieses MPIB-Projektes ist vom MPIB noch nicht veröffentlicht worden. (Das wurde bei der Verwendung des Begriffes *„Sekundäranalyse“* offenbar nicht berücksichtigt.)

Im Jahre 2002 hatten die Professoren Olaf Köller und Jürgen Baumert in dem von Rolf Oerter und Leo Montada neu herausgegebenen Handbuch der *„Entwicklungspsychologie“* anhand dieser älteren, beim Max-Planck-Institut für Bildungsforschung (MPIB) vorliegenden Daten die Vorzüge der in Deutschland üblichen frühen Differenzierung und - im Zusammenhang damit - den niedrigen Fördereffekt sechsjähriger Grundschulen beschrieben. Sie beriefen sich dabei auf die 1991 bzw. 1997 veröffentlichten Ergebnisse der erwähnten Studie von Peter M. Roeder und Fritz Sang. Die Daten dazu waren im Rahmen des MPIB-Projektes *„Schulleistung“* eingeholt worden.

Grundlage der Mitteilung von Roeder und Sang bzw. der Mitteilungen von Köller und Baumert waren *„verarbeitbare Datensätze von 12.594 Schülern“* aus 427 Gymnasien der damaligen Bundesrepublik und aus West-Berlin, mit Angaben zu den am Anfang und am Ende des 7. Jahrgangs in den Fächern Deutsch, Englisch und Mathematik vorgefundenen Leistungen (Baumert/Roeder/Sang/Schmitz 1986, S.644).

Unter der Überschrift *„Frühe Differenzierung fördert leistungsstarke Schüler“* referieren Köller und Baumert im Oerter/Montada 2002 folgende auf Daten des MPIB-Projektes *„Schulleistung“* (1968-1970) basierende Erkenntnis: ***„Unübersehbar zeigen sich deutliche Leistungsvorteile der Gymnasiasten aus Bundesländern mit vierjähriger Grundschule.*** *Die frühe Differenzierung scheint also leistungsstärkere Schüler in allen drei Fächern zu fördern.“ - „Zusammenfassend zeigt sich zumindest für das deutsche Schulsystem, dass bezogen auf die Fachleistungsentwicklung leistungsstarke Schüler von der Differenzierung im Sekundarbereich profitieren.“* (S.770/771)

Diese auf der Datengrundlage des MPIB-Projektes „Schulleistung“ beruhenden Erkenntnisse haben offenbar eine *„systematische“*, also nicht zeitgebundene Bedeutung. Sie sind daher für Köller und Baumert nach wie vor gültig. Denn sie werden bei gleicher Gelegenheit anhand einer im Jahre 2001 von O. Köller und A. A. Karim durchgeführten Auswertung von Daten des MPIB-Projektes *„Bildungsverläufe und psychosoziale Entwicklung im Jugendalter* (BIJU)*“* (1991-2001) aktualisiert und

sogar dahingehend ergänzt, *„dass bei Kontrolle der Ausgangsintelligenz und der familiären Herkunft die Intelligenzentwicklung am Gymnasium deutlich günstiger war.“ - „Mit dem günstigeren Lernklima gehen offenbar nicht nur günstigere Schulleistungsentwicklungen einher, sondern auch eine besondere Ausschöpfung kognitiver Begabungsreserven.“* (Oerter/Montada 2002, S.778)

Offenbar waren diese wichtigen Informationen des MPIB zum niedrigen Fördereffekt eines über den 4. Jahrgang hinausgehenden Unterrichts in undifferenzierten Klassen beim Kultusministerium von Mecklenburg-Vorpommern im Jahre 2005 noch nicht angekommen. Denn *„Regionalschulen“, „Gemeinschaftsschulen“, „Sekundarschulen“, „Allgemeine Sekundarschulen“* und *„schulart-unabhängige Orientierungsstufen“* arbeiten auch in den 5. und 6. Jahrgängen noch mit undifferenzierten Klassen.

Als am 27.02.2002 im Landtag von Mecklenburg-Vorpommern über die Einführung der Regionalschule beraten wurde, war ich zur Anhörung eingeladen. Nach meinem Vortrag musste ich mit Erstaunen feststellen, dass den Ministerialbeamten, die dort die Einführung der Regionalschule betrieben, das MPIB-Projekt „BIJU“ und seine Ergebnisse nicht bekannt waren, obgleich Mecklenburg-Vorpommern zu den vier am Projekt „BIJU“ beteiligten Bundesländern gehörte.

Aufgrund von Daten zum *„Regionalen Intelligenz-Durchschnitt* (RID)*“* der einzelnen Bundesländer, die aus einer Bundeswehrstudie vorliegen, darf vermutet werden, dass Mecklenburg-Vorpommern unter den vier am MPIB-Projekt „BIJU“ beteiligten Ländern auch in jenen Jahren bereits das Schulsystem mit dem relativ höchsten Fördereffekt hatte - wie fünf Jahre später in PISA 2000-E unter den vierzehn Bundesländern, deren PISA-Daten ausgewertet werden konnten.

Die negativen Auswirkungen der in Mecklenburg-Vorpommern inzwischen eingeführten Veränderungen (Regionalschule und schulart-unabhängige Orientierungsstufe) können im Rahmen von PISA-Studien erst ab 2011 nachgewiesen werden (Näheres dazu in unserem Aufsatz *„PISA 2000...“*).

Die Einführung der schulart-unabhängigen Orientierungsstufe in Mecklenburg-Vorpommern hätte verhindert werden können. Allerdings haben Köller und Baumert im Jahre 2002 ihre Erkenntnisse zu den Vorzügen des mit dem 5. Jahrgang einsetzenden, gegliederten deutschen Schulwesens und zu den Nachteilen sechsjähriger Grundschulen nicht mit einer dem Aufwand und der Bedeutung des MPIB-Projektes *„Schulleistung“* entsprechenden Öffentlichkeitswirkung bekannt gemacht, sondern lediglich in der Fachliteratur, eben in der neuen Ausgabe des Oerter/Montada von 2002, veröffentlicht, wo diese Erkenntnisse unbeachtet und ohne Wirkung blieben. Im KMK-Bildungsbericht 2003 werden ihre Erkenntnisse nicht erwähnt.

Eine so überaus wichtige, *„bildungsrelevante“* Information hätte jedoch vom MPIB spätestens mit diesem von der Kultusministerkonferenz in Auftrag gegebenen Bildungsbericht 2003 in den Kapiteln *„Schülerkompetenzen im historischen Bereich“* (C2.3) oder *„Leistungszuwächse“* (C2.4) an die Entscheidungsträger in Landesregierungen und Kultusministerien weitergereicht werden müssen.

Professor Dr. Peter Martin Roeder war von 1973 bis zu seiner Emeritierung im Jahre 1995 Direktor am MPIB und dort als Vorgänger Baumerts von 1982 bis 1995 zuständig für den *„Forschungsbereich Schule und Unterricht“*.

In einem Gutachten für den Landtag von Sachsen-Anhalt (Kopie als Anlage), das mit dem Briefkopf des Institutes versehen war, hat Professor Roeder in seiner Eigenschaft als Direktor am MPIB im Jahre 1995 dringend davon abgeraten, in Sachsen-Anhalt die *„Förderstufe“* als eine schulart-unabhängige und mit undifferenzierten Klassen arbeitende Orientierungsstufe einzuführen: Im Grunde laufe dies nämlich auf eine Verlängerung der Grundschulzeit von vier auf sechs Jahre hinaus und würde insofern leistungsstärkere Schüler absehbar benachteiligen.

Für dieses Gutachten berief sich Roeder auf Ergebnisse des MPIB-Projektes *„Schulleistung“* (1968-1970) und in Ergänzung dazu auch auf Ergebnisse des MPIB-Projektes *„Hauptschu-*

le/Gesamtschule“ (1980), welches gezeigt hatte, dass an weiterführenden Schulen die Probleme des Unterrichts in undifferenzierten Klassen durch Binnendifferenzierung nicht zu bewältigen sind.

Im Jahre 2005 hat es bei gleichem Anlass von Seiten des Max-Planck-Institutes für Bildungsforschung eine ähnliche Warnung an den Landtag von Mecklenburg-Vorpommern nicht gegeben, obgleich für Mecklenburg-Vorpommern aus dem MPIB-Projekt *„BIJU“* (1991-2001) beim MPIB sogar erheblich aktuellere Daten vorlagen.

Zwischenbemerkung: Es ist zweifellos eine der wesentlichen Aufgaben von „Bildungsberichten“, zur Beratung der Landesregierungen und der Kultusministerien stets den aktuellen Stand der Bildungsforschung zeitnah und in einer allgemein verständlichen Form zu referieren, damit Fehlentscheidungen - und die daraus entstehenden Fehlinvestitionen - verhindert werden können.

Im Bildungsbericht 2003 erfolgt im „Teil D“ (S.257-282) für die einzelnen Bundesländer auf 25 Seiten in großer Ausführlichkeit eine *„Systematisierte Darstellung zur Qualitätsentwicklung und zur Qualitätssicherung.“* Es fehlen jedoch zur beabsichtigten *„Erfassung des Zustands“* genauere Angaben zu den vorhandenen Stärken beziehungsweise zu den vorhandenen Schwächen der bestehenden Schulsysteme, es fehlt also eine *„Qualitätskontrolle des Vorhandenen“*.

Das wäre zum Beispiel möglich gewesen mit den „Qualitätskontrollen“ des Schulwesens von NRW und Berlin, Mecklenburg-Vorpommern und Sachsen-Anhalt, die am Max-Planck-Institut für Bildungsforschung (MPIB) aus dem MPIB-Projekt „BIJU“ bereits vorliegen, insbesondere mit einer genaueren Darstellung der sehr gründlichen und vom MPIB bereits veröffentlichten Qualitäts-Kontrollen des Schulwesens von Nordrhein-Westfalen aus dem MPIB-Projekt „BIJU“ (s.u.).

Der „Teil C“ des Bildungsberichtes 2003 enthält in den Kapiteln C2.3 (*„Schülerkompetenzen im historischen Vergleich“*) und C2.4 (*„Leistungszuwächse“*) einige Angaben zu diesen bereits vorhandenen Qualitätskontrollen des Max-Planck-Institutes für Bildungsforschung. Aber diese Angaben sind viel zu knapp, nicht sehr gründlich, unvollständig und bezüglich des **NRW-BIJU-Oberstufen-Vergleichs Mathematik von 1997** (S.200/201) irreführend missverständlich (s.u.). Die Ergebnisse des MPIB-Projektes *„Schulleistung“* (1968-1970) sind nicht jener Bedeutung entsprechend beschrieben worden, die sie für Köller und Baumert offenbar immer noch haben.

Es wären außerdem im KMK-Bildungsbericht 2003 innerhalb des Kapitels *„Leistungszuwächse“* (Teil C2.4, S.198-201) die von Jürgen Baumert und Olaf Köller in Pädagogik 6/1998 (S.17) veröffentlichten Ergebnisse des **NRW-BIJU-Mittelstufen-Vergleichs für Mathematik von 1995** (s.u.) ein wirklich orientierender Beitrag zur Schulform-Debatte gewesen. Und sie wären für den fachfremden Leser sehr viel verständlicher gewesen als die dort (S.199) veröffentlichte, aus dem gleichen Jahre stammende *„Tabelle C2/3“* von Olaf Köller.

Beide Bildungsberichte enthalten sehr viele Statistiken und Grafiken, die offenkundig mit großer Sorgfalt zusammengestellt worden sind, unter anderem etwa auch zur unterschiedlichen finanziellen und arbeitsmarktpolitischen Situation in den einzelnen Bundesländern. Aber es gibt für die Bundesländer keine Informationen zum unterschiedlichen Fördereffekt von Schulformen und Schulsystemen.

Es darf also weiterhin ungehindert die Einführung sechsjähriger Grundschulen gefordert werden, obgleich bezüglich dieser Forderungen durch den Aufsatz *„Über die institutionelle Verarbeitung von Leistungsunterschieden“*, der im Jahre 1991 von Baumerts Vorgänger Peter M. Roeder und Fritz Sang veröffentlicht wurde, eigentlich längst nachgewiesen wurde, dass in den 5. und 6. Jahrgängen von sechsjährigen deutschen Grundschulen leistungsstärkere Schüler nicht ihrer Begabung entsprechend gefördert werden. Auf eben diese Veröffentlichung berufen sich Köller und Baumert ja in ihrem bereits erwähnten Beitrag zum Oerter/Montada 2002 über die Vorteile einer mit dem 5. Jahrgang einsetzenden *„frühen Differenzierung“* und über die Nachteile von sechsjährigen Grundschulen.

Fazit: In der ***„Neuen deutschen Schulform-Debatte“*** fehlen *„empirisch belastbare“*, datengestützte, aktuelle Auskünfte über den günstigsten Zeitpunkt des Übergangs in die weiterführenden Schulen.

Soviel ist allerdings jetzt schon erkennbar: Es geht auch hier nicht ohne Kompromisse. Schon der „Weimarer Schulkompromiss“ von 1919/20 war eine solche Übereinkunft (und wie es scheint, eine gute). Es handelt sich eben um ein *Optimierungsproblem*, weil die beiden Zielvorgaben *„Soziale Integration“* und *„Individuelle Förderung“*, wie erwiesen, nicht gleichzeitig optimal zu verwirklichen sind.

Daher muss ein *„datengestützter“* Mittelweg gefunden werden, damit überprüfbare Argumente an die Stelle von gutgemeinten, aber agitativ missbrauchbaren Behauptungen treten.

3. Es fehlt zur Zeit eine aktuelle, datengestützte Diskussion über die bereits 1984 von Helmut Fend beschriebenen Vorteile einer Dreier-Differenzierung - gegenüber der zur Zeit in manchen Bundesländern diskutierten Zweier-Differenzierung.

Noch ein weiteres Thema der Schulformdebatte bedarf dringend der Klärung durch *„datengestützte Analysen“*: Helmut Fend hat bereits 1984 in seinem Aufsatz *„Determinanten von Lernleistungen: Wie wichtig sind die Lehrer?“* darauf hingewiesen, dass die Lernleistungen bei einer Dreier-Differenzierung des Unterrichts deutlich höher sind als bei einer Zweier-Differenzierung. Wir zitieren im Folgenden die Zusammenfassung seines Aufsatzes von 1984 (S.68):

„Die neueren amerikanischen Untersuchungen haben zu einer Rehabilitierung von Schule und Lehrern als wichtigen Determinanten von Schulleistungen geführt. Hier wird auf der Grundlage eines Teiles des größten deutschen Datensatzes zu Schulleistungen der Hypothese nachgegangen, dass die große Bedeutung des Lehrers nur dann sichtbar wird, wenn eine geringe institutionelle Vorstrukturierung von Schulleistungsniveaus vorliegt. So scheint es zu sein, da man bei hoch strukturierenden Schulsystemen durch zwei Informationen (Durchschnitt der Intelligenz einer Klasse und der sozialen Schicht) das Leistungsniveau einer Schulklasse in hohem Maße vorhersagen kann. Haben Lehrer institutionell mehr Freiheit, dann werden ihre persönlichen Haltungen zum Guten und Schlechten bedeutsamer.“

Helmut Fend war von 1969 bis 1982 der Projektleiter der wissenschaftlichen Begleitung des Gesamtschulversuchs *(„Konstanzer Studien“)*. Anhand der ihm zur Verfügung stehenden großen Datensätze aus Hessen, Niedersachsen und Nordrhein-Westfalen hat er mit *„Vergleichen vergleichbarer Lerngruppen“* nachgewiesen, dass die im traditionellen Schulwesen und an hessischen Gesamtschulen übliche Dreier-Differenzierung zu deutlich höheren Leistungen führt als die an den Gesamtschulen von Nordrhein-Westfalen und Niedersachsen übliche Zweier-Differenzierung. (Wie wir von Fend erfuhren, gehörte dieser Aufsatz zu den Vorarbeiten für seinen „Soester Vortrag“ von 1987. Bei diesem Vortrag distanzierte er sich mit ausführlicher Begründung von einigen Prämissen und einigen Hoffnungen der Gesamtschul-Initiatoren.)

Grundlage von Fends Untersuchung war die Leistungsentwicklung *„vergleichbarer Lerngruppen“* im Fach Englisch, und zwar von 67 Lerngruppen des dreigliedrigen Schulwesens, von 44 Lerngruppen aus dreier-differenzierenden hessischen Gesamtschulen und von 33 Lerngruppen aus den die Zweier-Differenzierung praktizierenden Gesamtschulen Nordrhein-Westfalens und Niedersachsens.

Das Ergebnis: *„In den Englischleistungen unterscheiden sich das herkömmliche Bildungswesen und das dreierdifferenzierende Gesamtschulsystem nicht. Das zweierdifferenzierende Gesamtschulsystem in Nordrhein-Westfalen und Niedersachsen tut dies allerdings in deutlichem Maße.“* - *„In der Gesamtschulvariante der Zweierdifferenzierung haben wir es global mit geringeren Leistungen, höherer Schülerzuwendung und geringerer Kontrolle und Strukturierung zu tun.“* (1984, S.79/80)

In diesem Zusammenhang macht Fend eine Mitteilung, die manche Behauptungen der Einheitsschul-Propaganda sehr in Frage stellt:

„Möglicherweise ist das deutsche Schulsystem in höherem Maße egalisierend, da es Schüler innerhalb stabiler Lerngruppen langfristig und konsequent fördern kann. Dass sich dabei Lehrer gegenüber Kindern aus verschiedenen sozialen Schichten bewusst diskriminierend verhalten, ist unwahrscheinlich.“ (1984, S.71; nachträglich unterstrichen)

Es gibt unübersehbare, gewichtige Gründe zu der Annahme, dass mit den Daten aus PISA 2000-E und aus PISA 2003-E die Vorteile eine Dreier-Differenzierung darzustellen sind (s. PISA 2000/01, S.128). In diesem Zusammenhang verweisen wir auf unseren Aufsatz *„Die Realschule im Spiegel der Bildungsforschung“* und auf die dort vorgelegten Begründungen. Zumindest aus PISA 2000-E liegen für die untersuchten Schüler neben den Daten zur sozialen Herkunft auch Daten zur kognitiven Grundfähigkeit sowie zu der ebenfalls *„vorhersagemächtigen Variable“* der Dekodierfähigkeit vor.

Anhand dieser Daten kann nachgewiesen werden, dass in Deutschland dreier-differenzierende Schulsysteme erheblich effektiver sind als Schulsysteme mit Zweier-Differenzierung. Und aus PISA 2003 liegen jetzt auch noch Daten vor zu einer dritten *„vorhersagemächtigen“* und als *„Prädikator“* verwendbaren Variable: zur *„Problemlöse-Kompetenz“*.

Fazit: In der ***„Neuen deutschen Schulformdebatte“*** fehlen aktuelle *„datengestützte“* Informationen zu dem Sachverhalt, dass die tradierte Dreier-Differenzierung höhere Lernleistungen möglich macht als die zur Zeit diskutierte Zweier-Differenzierung. Eine Zweier-Differenzierung (Y-Modell) ist in zwei Bundesländern in den 5. und 6. Jahrgängen bereits eingeführt worden. In zwei weiteren Bundesländern wird die Einführung eines zweigliedrigen Schulwesens von zwei Parteien und zwei Lehrerverbänden mit Eifer propagiert**.**

4. Es fehlen zur Zeit Berichte über die Ergebnisse der Vergleiche von Schulformen und Schulsystemen, die nach den *„Grundregeln der Schuleffektivitäts-Forschung“* als *„Vergleiche vergleichbarer Schüler auf drei unterschiedlichen, schulform-bezogenen Vergleichs-Ebenen“* durchgeführt worden sind.

Generelle Vorgabe: Nach den *„Grundregeln der Schuleffektivitätsforschung“* sind Qualitäts-Kontrollen im Bildungswesen erst dann relativ nah an der Realität, wenn nicht nur *„die soziale Herkunft“* der Schüler (*das* Thema der späten 60er Jahre), sondern auch *„die kognitiven Grundfähigkeiten“* der Schüler berücksichtigt worden sind.

In beiden Bildungsberichten wird - ebenso wie in den fünf bisher veröffentlichten Bildungsberichten der PISA-Studien - sehr viel über ***„sekundäre, sozialschicht-bedingte Disparitäten“*** referiert. Aber keiner dieser sieben Berichte enthält Angaben zu den Ursachen der ***„primären, begabungsbedingten Disparitäten“***, die am Ende der Grundschule für den Übergang auf eine der drei weiterführenden Schulformen von hoher Bedeutung sind, nämlich Angaben zu den kognitiven Grundfähigkeiten der untersuchten Schüler.

Wenn das mit den Daten aus den nationalen Ergänzungsstudien PISA 2000-E und PISA 2003-E nicht möglich sein sollte, wäre das unseres Erachtens eine Lücke im Konzept dieser sehr aufwändigen, kosten-intensiven Untersuchungen.

Keiner der beiden Bildungsberichte und keiner der bisherigen fünf PISA-Berichtsbände enthält *„Vergleiche vergleichbarer Schüler, die auf drei verschiedenen schulform-bezogenen Vergleichsebenen durchgeführt worden sind“*. Es fehlen daher Angaben zu den an den weiterführenden Schulen entstehenden, *„schulform-bedingten Disparitäten“*. Die müssten dann als ***„tertiäre Disparitäten“*** bezeichnet werden.

Auskünfte über die *tertiären Disparitäten* sind darum besonders wichtig, weil eigentlich nur diese Disparitäten ohne größeren Aufwand zum Guten hin verändert werden können: Durch Verbesserung der Effizienz von Schulformen oder durch Veränderung der Schulstruktur eines Landes. Begabung und Herkunft sind hingegen vorgefundene, relativ stabile und insofern nur noch begrenzt in ihren Auswirkungen zu beeinflussende *„Determinanten von Leistung in der Schule“*.

Dass inzwischen beim MPIB aufbereitete Datensätze aus drei großen Leistungsstudien vorliegen, die neben den Informationen zur *sozialen Herkunft* und zum *familiären Hintergrund* der Schüler auch Angaben zur *kognitiven Grundfähigkeit* der Schüler enthalten, ist in der Öffentlichkeit nicht hinreichend bekannt.

In den Daten-Speichern des MPIB lagern Daten aus dem MPIB-Projekt *„Schulleistung“* (1968-1970). Sie enthalten Angaben zur kognitiven Grundfähigkeit und zum familiären Hintergrund von mehr als 12.000 Gymnasiasten des 7. Jahrgangs aus den zehn Ländern der damaligen Bundesrepublik und aus West-Berlin. Zur Bestimmung des familiären Hintergrundes war nach Bildungsstand und Berufsposition der Eltern- und der Großelterngeneration gefragt worden (Edelstein 1970, S.526).

Peter M. Roeder und Fritz Sang haben 1991 in ihrem Aufsatz *„Über die institutionelle Verarbeitung von Leistungsunterschieden“* (1991, S.166-168) anhand dieser Daten durch den Vergleich der Leistungsentwicklung von *„gleich befähigten Schülern“* den niedrigen Fördereffekt der sechsjährigen Berliner Grundschulen nachweisen können. Auf die Daten der dort (1991, S.166) veröffentlichten *„Tabelle 1“* beriefen sich, wie erwähnt, Olaf Köller und Jürgen Baumert in ihrem Beitrag zum Oerter/Montada 2002 (S.770/771) über die Vorteile einer *„frühen Differenzierung“* und den niedrigen Fördereffekt sechsjähriger Grundschulen.

In den Daten-Speichern des MPIB lagern die aufbereiteten Datensätze des 1991 in vier Bundesländern im 7. Jahrgang aller Schulformen gestarteten MPIB-Projektes *„Bildungsverläufe und psychosoziale Entwicklung im Jugendalter* (BIJU)*“* (1991-2001). Sie enthalten Angaben zum familiären Hintergrund der Schüler sowie zu der am Anfang des 7. und am Ende des 10. Jahrgangs vorgefunden kognitiven Grundfähigkeit der an der BIJU-Längsschnittanalyse beteiligten etwa 9.000 Schüler aus NRW, BE, MV, ST. Im MPIB lagern auch die Ergebnisse einer in den Jahren 2000/2001 durchgeführten postalischen Befragung von etwa 3.000 BIJU-Teilnehmern über ihren weiteren Lebensweg. Von dieser Befragung erwartete man Auskünfte darüber, wie die jungen Leute rückblickend ihre schulische Ausbildung beurteilen.

In den Daten- Speichern des MPIB lagern außerdem auch aufbereitete Datensätze von mehr als 30.000 Neuntklässlern aller Schulformen aus allen Ländern der Bundesrepublik für PISA 2000-E, dessen Federführung beim MPIB lag, und zwar ebenfalls mit Angaben zum familiären Hintergrund und zur kognitiven Grundfähigkeit.

Bemerkungen zur Datenbasis für PISA-Schulform-Vergleiche

„Die kognitive Grundfähigkeit der Schülerinnen und Schüler, als relativ interventionsfernes Merkmal, wurde mit verbalen und figuralen Subtests erhoben, die Komponenten abstrakter räumlicher und verbaler Intelligenz abbilden. Zur Abschätzung der Effekte von anderen Variablen ist es sinnvoll, Unterschiede zwischen den Jugendlichen hinsichtlich ihrer kognitiven Grundfähigkeit konstant zu halten und damit zu prüfen, inwiefern andere Schülermerkmale einen zusätzlichen Beitrag zur Erklärung von Leistungsunterschieden leisten. Eine ebenfalls vorhersagemächtige Variable ist die Dekodierfähigkeit der Schülerinnen und Schüler, die sich in der Schnelligkeit des Erfassens der korrekten Bedeutung von Sätzen eines längeren Textes äußert. “ (PISA 2000/01, S.128; nachträglich unterstrichen) Auch für die unterschiedliche Fähigkeit des Dekodierens liegen beim MPIB aus PISA 2000-E entsprechende Daten vor.

Zwischenbilanz: Mit diesen Daten aus PISA 2000-E und aus PISA 2003-E wären also realitätsnahe Vergleiche von Schulformen und Schulsystemen möglich. Die hier vom MPIB angedeuteten Möglichkeiten einer Effizienz-Analyse mit PISA-Daten sind im Rahmen der bisherigen deutschen PISA-Berichte noch nicht für genauere Schulform-Vergleiche angewandt worden und daher in der deutschen Diskussion der PISA-Befunde auch noch nicht wahrgenommen worden.

Genauere Schulformvergleiche sind erst jene Schulformvergleiche, die auf den drei verschiedenen schulform-bezogenen Vergleichs-Ebenen durchgeführt worden sind (s.u.).

Wie F. E. Weinert 1999 bei der Hauptversammlung der Max-Planck-Gesellschaft vorgetragen hat, sind die kognitiven Grundfähigkeiten die wichtigsten Determinanten schulischer Leistungen.

Aber dieses Thema ist seit der von Heinrich Roth betriebenen Propagation eines *„dynamischen Begabungsbegriffes“* das erklärte „Tabu“ der deutschen Bildungsforschung.

Mit den Daten aus PISA 2003-E zur *„Problemlöse-Kompetenz“* der untersuchten Schüler liegen Angaben zu einer dritten *„vorhersagemächtigen Variable“* vor. Diese drei *„vorhersagemächtigen Variablen“* können als *„Prädikatoren“* genutzt werden für eine *„prädikatorengestützte Analyse“* der vorgefundenen Leistungen.

Mit Hilfe dieser drei Prädikatoren ist dann auch anhand von PISA-Befunden festzustellen, ob die Leistungen der Schüler *„erwartungsgemäß“* oder *„erwartungswidrig negativ“* oder *„erwartungswidrig positiv“* ausgefallen sind (vgl. PISA 2000/03, Baumert, Trautwein, Artelt, S.286f).

Durch die Beschränkung auf Neuntklässler ist eine höhere Realitätsnähe der Befunde möglich, weil dann noch genauere Leistungsvergleiche möglich sind. Es bleiben zum Beispiel jene Schüler außen vor, die zu spät eingeschult worden oder irgendwann sitzengeblieben sind, also alle Fünfzehnjährigen mit dem Leistungsstand des 8. Jahrgangs.

Michael Neubrand und Eckhard Klieme begründen die Begrenzung auf Neuntklässler folgendermaßen (PISA 2000/02, S.114; die Unterstreichung wurde nachträglich eingefügt):

„Da die Jahrgangsklasse in den Schulsystemen aller Länder der Bundesrepublik Deutschland bis heute die maßgebliche organisatorische Einheit geblieben ist und da die Leistungsanforderungen der Schule auf Klassenstufen bezogen definiert werden, gibt der Vergleich der Neuntklässler schulsystemnah Auskunft über die Leistungen der Schülerinnen und Schüler in den einzelnen Ländern.“

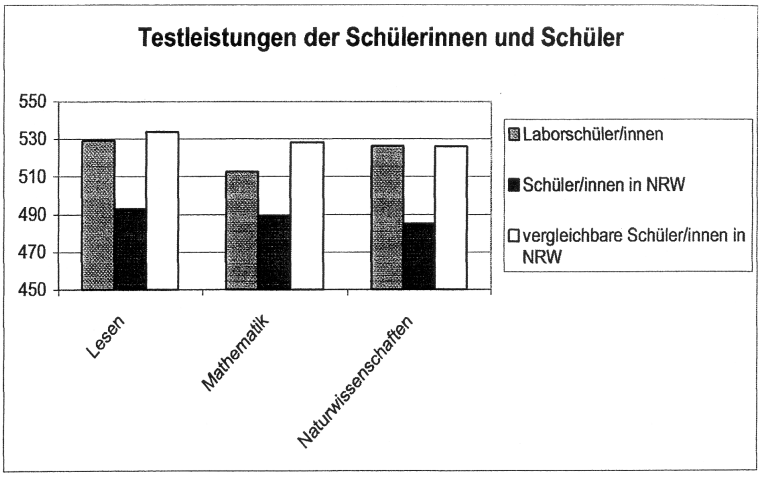
Aus diesem Grunde gab es für PISA 2000-E neben den Ranglisten der gesamten Stichprobe auch Ranglisten der Neuntklässler. Die beiden ersten Berichtsbände aus PISA 2003 enthalten jedoch seltsamerweise keine Ranglisten von Neuntklässlern. Wenn auch im dritten Berichtsband von PISA 2003-E keine Vergleiche von Neuntklässlern vorgelegt werden, dann verstieße die Auswertung der Daten aus PISA 2003-E gegen jene von Michael Neubrand und Eckhard Klieme vorgetragenen prinzipiellen Erwägungen.

Fazit: Mit den am MPIB aus drei großen Leistungsstudien vorliegenden Daten zur kognitiven Grundfähigkeit könnten für das deutsche Schulwesen endlich auch die unbestreitbaren schulform-bedingten tertiären Disparitäten genauer beschrieben werden.

Denn zumindest mit den Daten aus PISA-2000-E sind Vergleiche vergleichbarer Schüler möglich, bei Anwendung der von Baumert und Köller 1998 (Pädagogik 6/1998, S.17) beschriebenen Methode sogar auf drei verschiedenen, schulform-bezogenen Vergleichs-Ebenen.

Das entsprechende Verfahren ist im dritten Band der Veröffentlichungen aus PISA 2000/03 (S.286) von Jürgen Baumert und seinen Mitarbeitern ausführlich erläutert und am Beispiel der Laborschule Bielefeld eindrucksvoll vorgeführt worden.

Nachfolgend eine Grafik aus dem Anhang der Presse-Information, die vom MPIB bei der Pressekonferenz der Laborschule Bielefeld am 13.11.2002 vorgelegt wurde, um die Ergebnisse eines mit PISA-Daten durchgeführten Leistungsvergleichs dieser Schüler zu veranschaulichen. Titel der Presseinformation des MPIB: *„Kompetenzerwerb und Persönlichkeitsentwicklung: Eine Untersuchung an der Laborschule Bielefeld im Rahmen von PISA“*. (Der Ausdruck *„vergleichbare Schüler“* ist von uns aus dem Text dieser Presseinformation des MPIB übernommen worden.)



Kommentar jener MPIB-Presse-Information vom 13.11.2002 zu nebenstehender Grafik: *„Dieser Vergleich zeigt, dass in den Bereichen Lesen und Naturwissenschaften die Leistungen der 15-Jährigen ungefähr den Leistungen vergleichbarer Schülerinnen und Schüler anderer Schulen entsprechen. In Mathematik bleiben die Leistungen der Laborschülerinnen und Laborschüler <u>etwas</u> unter dem Mittelwert der Vergleichsgruppe.“* (S.2; Unterstreichung nachträglich eingefügt.) Mit etwa 15 Punkten der PISA-Skalierung hatten die Schüler dieser hochbegünstigten Gesamtschule in Mathematik innerhalb der Vergleichsgruppe einen Leistungsrückstand von fast einem halben Schuljahr, das ist etwas mehr als *„etwas“*! Bei einem Mittelwert M = 500 entsprechen 30 bis 35 Testpunkte in Mathematik dem Lernfortschritt eines ganzen Schuljahres (vgl. Köller u.a. 2004, S.691).

Es ist dieser Vergleich - unseres Wissens - bisher der einzige vom MPIB mit PISA-Daten durchgeführte Vergleich zur unterschiedlichen Leistungsentwicklung *„vergleichbarer Schüler“* von weiterführenden Schulen, allerdings noch nicht auf drei verschiedenen, schulform-bezogenen Vergleich-Ebenen. Aber es ist der Präzedenzfall einer derartigen Verwendung von PISA-Daten! Es kann also nicht behauptet werden, mit PISA-Daten seien derartige *„approximative“* und insofern verlässlich orientierende Schulform-Vergleiche nicht möglich (vgl. KMK-Bildungsbericht 2003, S.203).

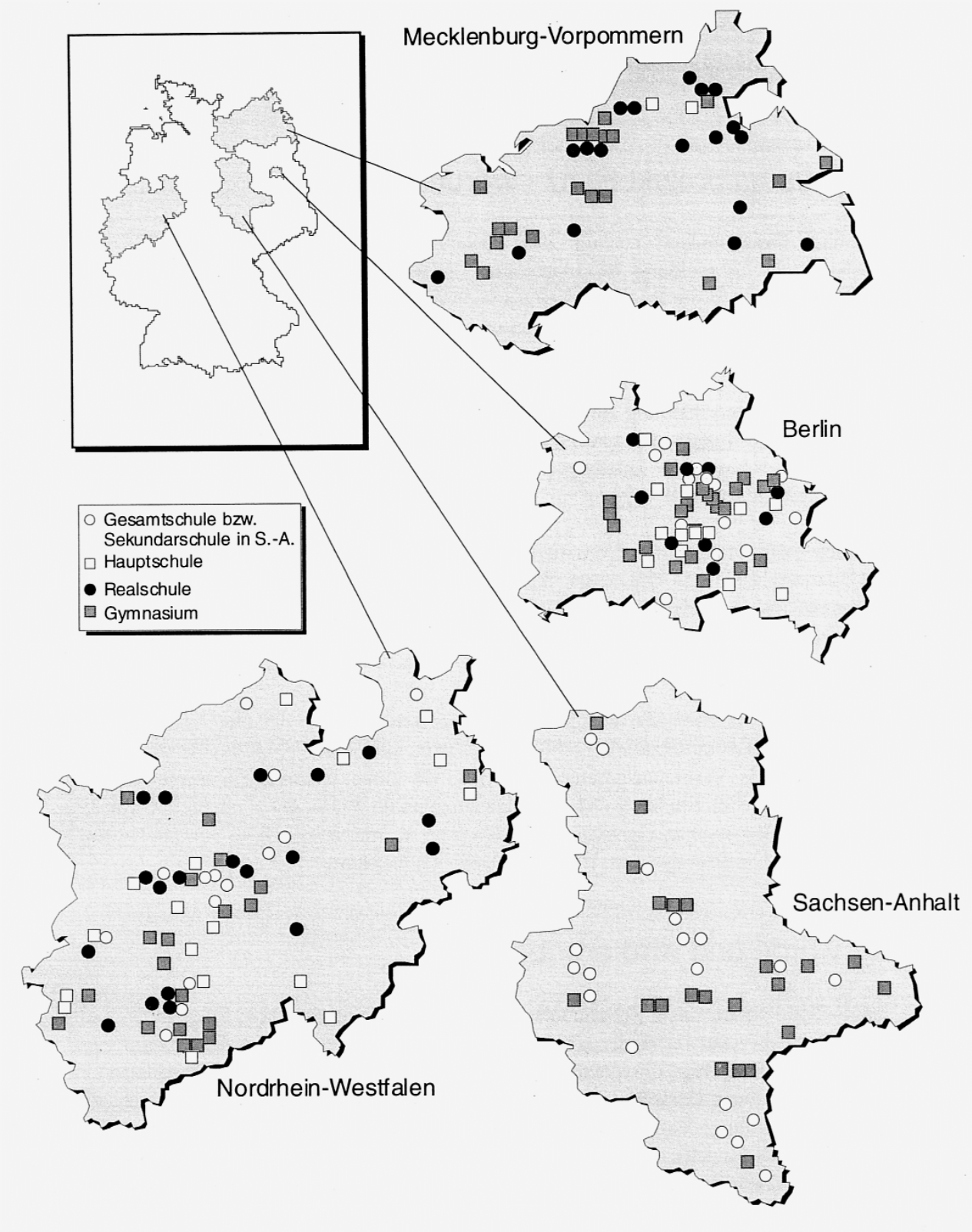
Denn wenn durch den Vergleich von Laborschülern mit *„vergleichbaren Schülerinnen und Schülern in NRW“* Erkenntnisse über den Fördereffekt der Laborschule Bielefeld gewonnen werden konnten, dann können doch ebenso auch *„vergleichbare Gruppen“* dieser anderen NRW-Schüler <u>untereinander</u> verglichen werden, um so die unterschiedlichen Fördereffekte von <u>deren</u> Schulen zu ermitteln.

Bei solchen Vergleichen geht man davon aus, dass sich in den meisten Fällen *„die sozialen Ausgangsbedingungen“* im Laufe der Schulzeit nicht wesentlich verändern. Die bei den PISA-Studien im 9. Jahrgang vorgefundenen Leistungen von *vergleichbaren* Schülern geben dann in Annäherung Auskunft darüber, wie <u>unterschiedlich</u> diese *vergleichbaren* Schülerinnen und Schüler an den verschiedenen Schulformen und Schulsystemen gefördert worden sind, - unterschiedlich trotz *„gleicher intellektueller und sozialer Ausgangsbedingungen“*.

Auch für PISA-2003-E sind Daten zur kognitiven Grundfähigkeit erhoben worden. Dies wurde uns bei einem vom Deutschen Philologenverband und vom Verband Deutscher Realschullehrer am 19. Oktober 2005 veranstalteten Symposium von einem Mitarbeiter des MPIB auf Anfrage bestätigt. Auch mit den Daten aus PISA 2003-E können Schulformen durch *„Vergleiche vergleichbarer Schüler“* als *„differenzielle Entwicklungsmilieus“* ausgewiesen werden. (DPhV und VDR: *Dokumentation: Die Zukunft des gegliederten Schulwesens*, 2006, S.52)

Standorte der Schulen, die am MPIB-Projekt-BIJU (1991 - 2001) teilgenommen haben

In Berlin waren es 13 Hauptschulen, 9 Realschulen, 23 Gymnasien und 14 integrierte Gesamtschulen.
In NRW waren es 17 Hauptschulen, 19 Realschulen, 21 Gymnasien und 14 integrierte Gesamtschulen.

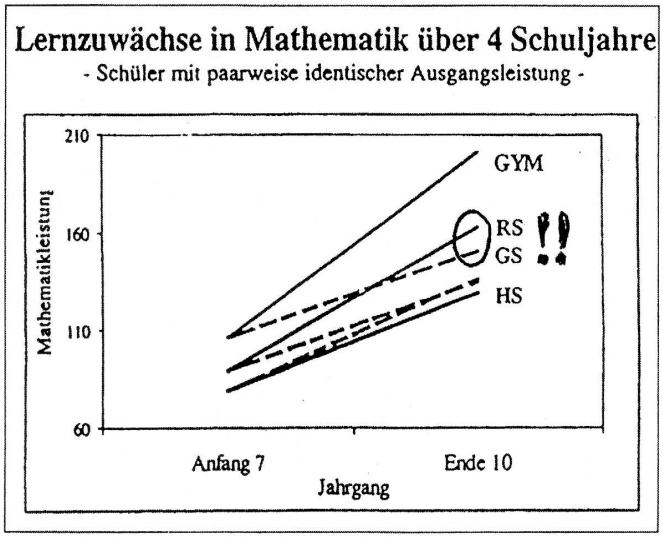


Quelle: Erster Zwischenbericht für die am Projekt „BIJU“ teilnehmenden Schulen, 1994, S.16, *Abbildung 1*

Um zu zeigen, welche bemerkenswerten Befunde eines Schulformvergleichs aus dem MPIB-Projekt *„Bildungsverläufe und psychosoziale Entwicklung im Jugendalter* (BIJU)“ in der Fachliteratur seit 1998 bereits vorliegen, nachfolgend als Beispiel eine ausführlichere Darstellung von Erkenntnissen zum unterschiedlichen Fördereffekt der verschiedenen Schulformen. Diese Befunde wurden mitgeteilt in einem Aufsatz, der von Baumert und Köller in der Zeitschrift „Pädagogik“ (6/1998, S.13-18) veröffentlicht worden ist. (Eine auf der Vorseite wiedergegebene Karte zeigt für alle vier Bundesländer die Standorte der Schulen, die am Projekt-BIJU teilgenommen haben.)

Es wurde bei diesem Vergleich auf drei schulform-bezogenen Vergleichsebenen jeweils eine Gruppe von Hauptschülern, von Realschülern und von Gymnasiasten einer *„vergleichbaren“* Gruppe von Gesamtschülern gegenübergestellt. Insgesamt waren es also sechs Gruppen *„mit paarweise identischen Ausgangsbedingungen“*. Die jeweils zu vergleichenden *Paare* hatten *ähnliche kognitive Grundfähigkeiten* und *einen ähnlichen familiären Hintergrund*. Sie hatten außerdem am Anfang des 7. Jahrgangs in Mathematik auch noch *den gleichen Leistungsstand*, in der weiter unten abgebildeten Grafik erkennbar am Start aus demselben Punkt. Die Schüler waren Schüler aus NRW.

Die hier abgedruckte Grafik wurde entwickelt nach einer in Farbe gehaltenen, bisher noch nicht veröffentlichten Grafik, die 1998 von Mitarbeitern des MPIB bei verschiedenen Informationsveranstaltungen vorgestellt worden ist. Sie veranschaulicht die in „Pädagogik“ 6/1998 (S.17) von Baumert und Köller folgendermaßen beschriebenen Befunde: *„Für den Vergleich von Haupt- und Gesamtschule ergeben sich nach Kontrolle des Vorwissens sowie der kognitiven und sozialen Variablen keine unterschiedlichen Leistungseffekte zwischen beiden Schulformen: Bei gleichen Eingangsbedingungen wird am Ende des 10. Jahrgangs ein identischer Wissensstand erreicht.“ - „Beim Vergleich zwischen Real- und Gesamtschule zeigt sich, dass in der Realschule auch nach Kontrolle kognitiver und sozialer Eingangsvariablen die Leistungsentwicklung günstiger als an der Gesamtschule verläuft: Bei gleichen intellektuellen und sozialen Eingangsbedingungen erreichen Realschüler am Ende der Sekundarstufe I etwa in Mathematik einen Wissensvorsprung von etwa zwei Schuljahren.“ - „Noch stärker sind diese Effekte, wenn man Gesamtschule und Gymnasium vergleicht: Bei gleichen intellektuellen und sozialen Bedingungen beträgt der Leistungsvorsprung in Mathematik am Gymnasium mehr als zwei Schuljahre.“ - „Es gibt keine Hinweise, dass die ungünstige Leistungsentwicklung durch besondere überfachliche Leistungen kompensiert werden könnte.“* (Pädagogik 6/1998, S.17; die Unterstreichungen wurden nachträglich eingefügt.)



Der Vergleich wurde seinerzeit durchgeführt und veröffentlicht, weil die von Köller 1996 im 2. BIJU-Bericht veröffentlichten kritischen Befunde zur Leistungsentwicklung an NRW-Gesamtschulen von Gesamtschulbefürwortern in recht aggressiver Form angezweifelt worden waren.

Zwischenbemerkung: In der ***„Neuen deutschen Schulform-Debatte“*** ist dieser von Baumert und Köller *auf drei verschiedenen schulform-bezogenen Vergleichs-Ebenen* durchgeführte Vergleich, mit dem erstmals der niedrige Fördereffekt der 5. bis 10. Jahrgänge von NRW-Gesamtschulen nachgewiesen wurde, überhaupt noch nicht zur Kenntnis genommen worden.

Unseres Wissens ist vom MPIB außer diesem auf drei verschiedenen schulform-bezogenen Vergleichs-Ebenen durchgeführten Vergleich nur noch ein einziger, ebenfalls auf einer schulform-bezogenen Vergleichs-Ebene durchgeführter Vergleich veröffentlicht worden.

Das ist der 1991 von Roeder und Sang als *„Tabelle 1“* veröffentlichte Vergleich zum unterschiedlichen Fördereffekt von vierjährigen und sechsjährigen Grundschulen. Mit den Regressions-Residuen dieses Vergleichs konnten Roeder und Sang nachweisen: Berliner und Bremer Gymnasiasten, die erst nach sechsjähriger Grundschule aufs Gymnasium übergewechselt waren, hatten am Anfang des 7. Jahrgangs gegenüber den Altersgenossen der übrigen Bundesländer, die dort schon nach vierjähriger Grundschule aufs Gymnasium überwechseln konnten, in Deutsch, Englisch und Mathematik einen Leistungsrückstand von mehr als einem Schuljahr. Auch ihre weitere Leistungsentwicklung war deutlich verzögert. Dies wurde anhand der „Regressions-Residuen“ nachgewiesen. Das sind die Ergebnisse eines Rechenverfahrens, bei dem die Auswirkungen der unterschiedlichen Vorkenntnisse, Begabungen und Milieus herausgerechnet und so die reine Förderleistung des Unterrichts dargestellt werden kann (vgl. Baumert/Roeder/Sang/Schmitz 1986, S.656, und Roeder/Sang 1991, S.164;166).

„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.“

„Die Aussagen der Lehrer über ihre Unterrichtsplanung für das 7. Schuljahr und für 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; nachträglich unterstrichen - Spätere Veröffentlichungen des MPIB hielten sich nicht immer an diese vom MPIB bereits 1968 beim MPIB-Projekt „Schulleistung“ eingeführten Standards.)

Es fehlen aus der Bildungsforschung insbesondere für das Fach Deutsch *„Vergleiche vergleichbarer Schüler“*. Denn nur so kann geklärt werden, ob Schüler eine Sache darum nicht verstehen, weil es ihnen an Verständnis für die Sache fehlt, oder lediglich darum, weil sie die Sprache, in der die Sache vorgetragen wird, nicht verstehen. Das gilt nicht nur für Schüler mit Migrations-Hintergrund.

Deutsch war bis 1997 ebenfalls Gegenstand des MPIB-Projektes „BIJU“. Im 1. Zwischenbericht aus dem Projekt „BIJU“ (1994) für die am Projekt „BIJU“ teilnehmenden Schulen werden an mehreren Stellen noch Untersuchungsergebnisse für das Fach Deutsch mitgeteilt. Die später vorgefundenen Leistungen waren anscheinend derart desaströs, dass sie nicht bekannt gemacht wurden.

Mit den Daten des MPIB-Projektes BIJU sind die von Roeder und Sang im Jahre 1991 berichteten Befunde zum niedrigen Fördereffekt der sechsjährigen Berliner Grundschulen *„repliziert“* worden. Aber das Ergebnis auch dieses Vergleichs ist vom MPIB nicht veröffentlicht worden (s. S.2 in:
http://www.mpib-berlin.mpg.de/de/forschung/eub/Entgegnung%20Sprenger%2025.03.03.pdf)

Alle fünf vom PISA-Konsortium Deutschland im Auftrag der KMK erstellten Berichtsbände enthalten keine Vergleiche von Schulformen und Schulsystemen, die nach Art des von Baumert und Köller 1998 veröffentlichten Vergleichs auf drei verschiedenen schulform-bezognen Vergleichs-Ebenen durchgeführt worden sind, bei denen also auch die kognitive Grundfähigkeit der untersuchten Schüler berücksichtigt worden wäre.

Und nun muss festgestellt werden, dass auch die beiden Bildungsberichte keine Mitteilungen über solche Vergleiche enthalten, obwohl mindestens zwei auf *Schulform-Ebenen durchgeführte Vergleiche vergleichbarer Schüler* vom MPIB bereits veröffentlicht wurden - und obwohl Professor Dr. Jürgen Baumert zum Beirat des Konsortiums gehört.

Fazit: Solange bei einer auf Leistungsentwicklung angelegten Zustandsbeschreibung des Bildungswesens die kognitiven Grundfähigkeiten der Schülerschaft nicht berücksichtigt werden, fehlen in den Bildungsberichten die Angaben zur wichtigsten Determinante schulischer Leistungen.

Für eine angemessene und faire Beurteilung der vorgefundenen Leistungsunterschiede sind sie dringend erforderlich.

5. Es fehlen zur Zeit für die einzelnen Bundesländer Angaben zum unterschiedlichen Begabungs-Potential ihrer Schülerschaft.

Nach den Grundregeln der Schuleffektivitäts-Forschung sind Qualitätskontrollen von Schulen, Schulformen und Schulsystemen erst dann relativ nah an der Realität, wenn nicht nur die *soziale Herkunft* der beteiligten Schüler, sondern auch ihre *kognitiven Grundfähigkeiten* berücksichtigt werden.

Wie Franz E. Weinert anhand von Ergebnissen der Zwillings-Forschung 1999 in seinem berühmten Festvortrag bei der Hauptversammlung der Max-Planck-Gesellschaft auf Grund jahrelanger Vergleichs-Studien der Zwillingsforschung überzeugend vorgetragen hat, sind Schulleistungen und daher auch die Unterschiede von Schulleistungen zu etwa 50% begabungsbedingt. Das gilt für bundesländer-interne ebenso wie auch für bundesländer-übergreifende Leistungsunterschiede.

Zwischenbilanz: In der ***„Neuen deutschen Schulform-Debatte“*** fehlen aus den PISA-Studien ebenso wie aus dem MPIB-Projekt „BIJU“ aktuelle, *„datengestützte“* Informationen über *die kognitiven Grundfähigkeiten* der untersuchten Schülerschaft. Die kognitiven Grundfähigkeiten sind die wichtigsten Determinanten von Leistung in der Schule. Denn die kognitiven Grundfähigkeiten sind die Hauptursache der in der gemeinsamen Grundschulzeit entstehenden *„primären, begabungs-bedingten Disparitäten“* und daher schon beim Übergang auf die weiterführenden Schulen von hoher Bedeutung.

Was hier über die Bedeutung der kognitiven Grundfähigkeiten bezüglich der Entstehung von Disparitäten vorgetragen wurde, gilt ebenso auch für eine adäquate Beurteilung der beim Vergleich der Bundesländer vorgefunden Leistungsunterschiede. Nur mit ihrer Hilfe sind die in der *„Einleitung“* des Bildungsberichtes 2006 (S.3) geforderten *„länderspezifisch vergleichbaren Aussagen“* möglich.

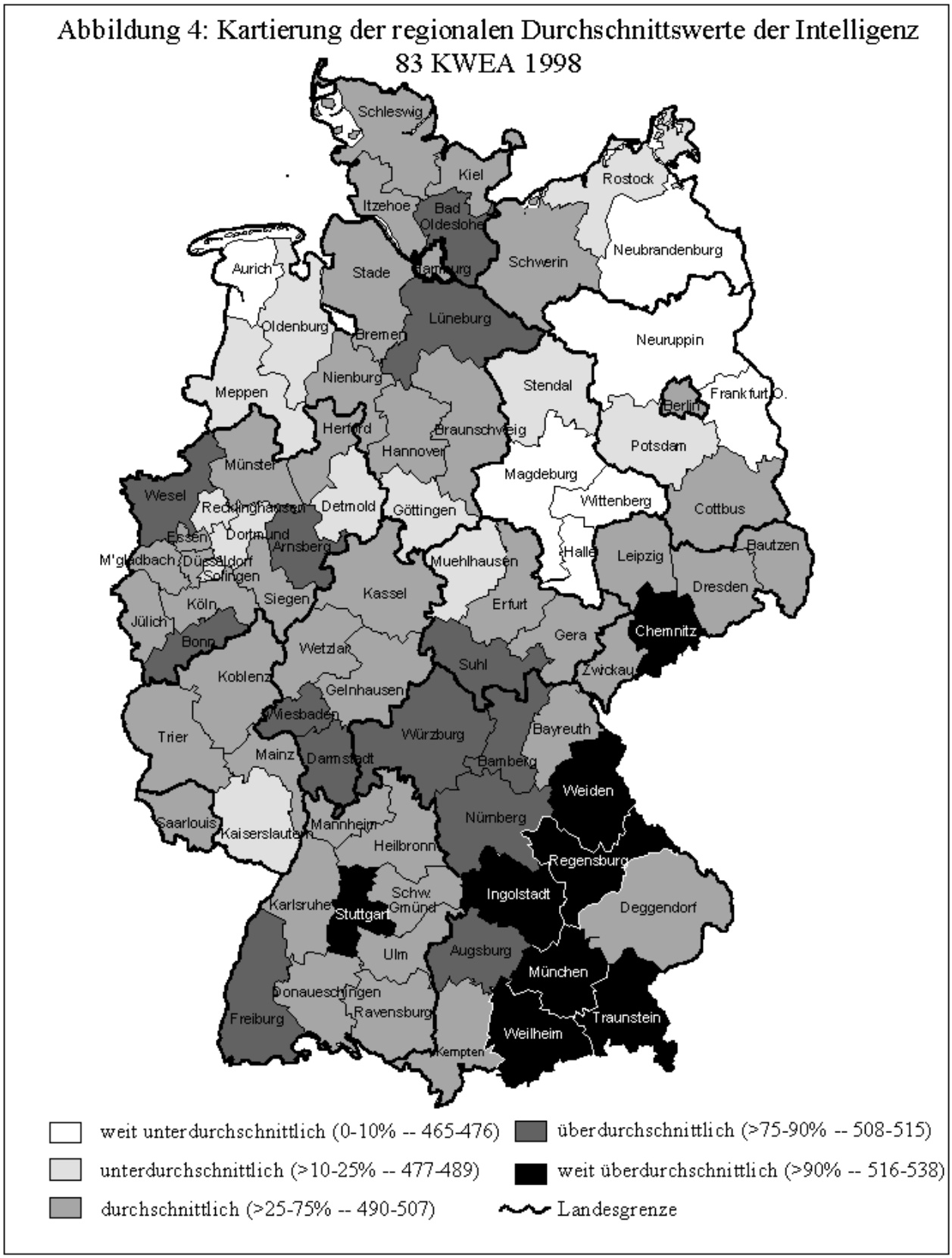
Denn erst wenn für die Schülerschaft eines Bundeslandes auch der *„Mittelwert ihrer kognitiven Grundfähigkeiten“* bekannt ist, können die Mittelwerte der vorgefundenen Leistungsunterschiede angemessen und fair beurteilt werden. (Im Zusammenhang mit diesen und weiteren Ausführungen verweisen wir auf unseren Aufsatz *„PISA 2000...“*.)

In den bisher veröffentlichten Ranglisten der Ländervergleiche aus PISA 2000-E und PISA 2003-E und in beiden KMK-Bildungsberichten ist nicht berücksichtigt worden, dass es zwischen den Ländern der Bundesrepublik offenbar Unterschiede beim Mittelwert des Intelligenz-Potentials ihrer Schülerschaft gibt, in der Fachsprache als *„Regionaler Intelligenz-Durchschnitt* (RID)*“* bezeichnet.

Weil derartige Angaben fehlen, ist zum Beispiel nicht zu erkennen, ob es nun an der Qualität der Schülerschaft oder an der Qualität des Schulsystems oder an der Qualität der Lehrerschaft liegt, dass die Leistungen der Bremer und Brandenburger Schüler so sind, wie sie sind.

Die wohl von Franz E. Weinert erstmals vorgelegte Frage: *„Schulleistungen - Leistungen der Schule oder des Schülers?“* ist jedenfalls in den bisherigen, mit großem Aufwand erstellten fünf PISA-Berichtsbänden nicht beantwortet worden - und in den Bildungsberichten von 2003 und 2006 auch nicht.

Erste orientierende Informationen zum *„Mittelwert des Intelligenzpotentials der einzelnen Bundesländer“* gibt es aus einer im Jahre 2002 veröffentlichten Studie der Bundeswehr, die auf den Tests der Musterung der etwa 250.000 Rekruten des Jahrgangs 1998 basiert.



Quelle: Ebenrett/Hansen/Puzicha (2002): „Brain drain“ in Deutschen Regionen, S. 11 -
Die Wiedergabe geschieht mit der freundlichen Genehmigung der drei Autoren.

Die Bedeutung dieser Kartierung ist klar, sobald man wahrgenommen hat, wie sehr sich Bremen, das Schlusslicht aller PISA-Ranglisten, hier als ein weißer Fleck von seiner Umgebung abhebt.

Die Studie der Bundeswehr beginnt mit folgendem Satz:

> *„Der Psychologische Dienst der Bundeswehr stellt beim Vergleich der Bundesländer seit langem auffällige Unterschiede in den Intelligenzleistungen wehrpflichtiger junger Männer fest. Markante Ost-West-Unterschiede stehen einem nicht minder starken Süd-Nord-Gefälle gegenüber.“*

Titel der Studie: *„Brain drain“ in deutschen Regionen - Effekte von Arbeitslosigkeit und innerdeutscher Migration“*. Ihre Verfasser sind Heinz J. Ebenrett, Dieter Hansen und Klaus J. Puzicha.

Es handelt sich bei dieser Untersuchung um eine im Juli 2002 als Manuskript veröffentlichte Studie des „Psychologischen Dienstes der Bundeswehr (PSZ III 6)“. Sie ist in überarbeiteter Fassung in der Beilage zur Wochenzeitschrift „Das Parlament“ vom 03.02.03 (S.25-31) veröffentlicht worden.

Diese Studie ist uns vom Psychologischen Dienst der Bundeswehr zusammen mit einer in der PISA-Skalierung gehaltenen RID-Rangliste der Bundesländer dankenswerterweise zur Verfügung gestellt worden. Beide Veröffentlichungen enthalten die hier abgebildete, kartierte Darstellung der Unterschiede des *„Regionalen Intelligenz-Durchschnitts* (RID)*“*. Ihre Verwendung geschieht mit der freundlichen Genehmigung der Autoren.

In der Kartierung sind die Bereiche der 83 Kreiswehrersatzämter (KWEA) noch einzeln dargestellt. Für die Rangliste der Bundesländer wurden diese Durchschnittswerte dann addiert und in die PISA-Skalierung (M=500, SD=100) umgerechnet.

Bayern liegt auch hier an der Spitze, und Bremen ist auch hier das Schlusslicht. Die Differenz liegt mit 520 zu 466 bei mehr als einer halben Standardabweichung. Das ist unseres Wissens weit mehr als der innerhalb eines Jahres zu erwartende Entwicklungsfortschritt der kognitiven Grundfähigkeiten bei Schülern dieses Alters.

Wie in der Bundeswehr-Studie festgestellt wurde, sind die Unterschiede im *„Regionalen Intelligenz-Durchschnitt* (RID)*“* überwiegend eine Folge der durch regionale Wirtschaftsprobleme verursachten *„innerdeutschen Migration“*: Tüchtige junge Leute wanderten ab in wirtschaftlich stärkere Regionen der Bundesrepublik. Die so entstandene Absenkung des *„Regionalen Intelligenz-Durchschnitts* (RID)*“* ist also ein *„externer Faktor“* von Leistungsunterschieden, die daher nicht dem Schulsystem angelastet werden sollten. (Unseres Erachtens dürfte wohl auch die 1945 einsetzende West-Flucht eine der Ursachen der Ost-West-Unterschiede sein.)

„Erst wenn die Bedeutung der Kontextbedingungen für die Schülerleistungen erfasst und kontrolliert ist, hat man eine Basis für <u>faire</u> Ländervergleiche, das heißt für Vergleiche, die <u>extern bedingte</u> Leistungsunterschiede zwischen den Ländern nicht ihrer bildungspolitischen und pädagogischen Praxis zuschreiben.“ (Jürgen Baumert und Manfred Weiß in PISA 2000/02, S.39, Unterstreichung nachträglich eingefügt) Ohne Zweifel gehören Unterschiede des *„Regionalen Intelligenz-Durchschnitts* (RID)*“* zu den *„externen Bedingungen“* von Leistungsunterschieden.

Alle bisherigen Veröffentlichungen aus den PISA-Studien geben keine Auskunft darüber, ob es nun an der Qualität des Schulsystems oder an der Qualität der Schülerschaft liegt, dass die vorgefundenen Leistungen so sind, wie sie sind. Daher bleiben die in weniger erfolgreichen Ländern tätigen Lehrerinnen und Lehrer dem Verdacht ausgesetzt, dass es am Ende nur an ihnen gelegen hat.

Wir halten es für eine kollegiale Pflicht, auf dieses ehrenrührige Informations-Defizit hinzuweisen. Denn vom MPIB ist wiederholt behauptet worden: *„Entscheidend für die Qualität von Unterrichtsprozessen ist bei jeder Organisationsform der pädagogisch intelligente Umgang mit heterogenen Leistungsvoraussetzungen von Schülern.“* So auch in einem von den Professoren Baumert, Köller und Roeder unterschriebenen Leserbrief des MPIB an DIE WELT vom 02.04.2003.

Rangliste der Bundeswehr-Studie bezüglich des Regionalen Intelligenz-Durchschnitts (RID)
- MW 500 / SD 100, gerundet -

520	Bayern
509	Hessen
508	Schleswig-Holstein
506	Baden-Württemberg
504	Sachsen
498	Thüringen
497	Hamburg
496	Rheinland-Pfalz
495	Nordrhein-Westfalen, Saarland
490	Niedersachsen
488	Berlin
481	Mecklenburg-Vorpommern
475	Brandenburg
469	Sachsen-Anhalt
466	Bremen

Quelle: Psychologischer Dienst der Bundeswehr PSZ III 6, 2002 - Basis: Eignungstests der etwa 250.000 Rekruten des Jahrgangs 1998

Neuntklässer aus PISA 2000: Gesamtsumme der Differenzen Leistungsmittelwerte minus RID

LK=Lesekompetenz, **M**=Mathematik, **NW**=Naturwissenschaften

Land		LK M NW
M-V:	**+29**	(+4/+16/+9)
S-A:	**+17**	(-3 /+13/+7)
B-W:	**+10**	(+4/ +6/ 0)
THÜ:	**+6**	(-2 / +2/ +6)
BAY:	**+3**	(-2 / +8/ -3)
SAC:	**+3**	(-3 / +4/ +2)
N-S:	**-2**	(+1/+3/ -6)
R-P:	**-3**	(0/ 0/ -3)
BRA:	**-10**	(-7/+2/ -5)
BRE:	**-13**	(-7/ -1/ -5)
SAA:	**-17**	(-8/ -3/ -6)
NRW:	**-25**	(-2/-11/-12)
S-H:	**-35**	(-18/ -5/-12)
HES:	**-68**	(-23/-20/-25)

Uns ist bekannt, dass auch am MPIB aus PISA 2000-E und aus dem MPIB-Projekt BIJU (1991-2001) ähnliche RID-Ranglisten vorliegen. Wenn nicht, wäre das ein schwerer Kunstfehler. Die Tatsache, dass wir das eine wissen, gibt uns die Freiheit, das andere zu behaupten. Uns ist außerdem bekannt, dass Herrn Professor Baumert bereits seit längerem die RID-Listen des Musterungs-Jahrganges 1998 vorliegen.

Bei den Ranglisten zum Regionalen Intelligenz-Durchschnitt, die vom MPIB erstellt worden sind, handelt es sich - im Gegensatz zu den RID-Listen der Bundeswehr - um Ranglisten für Jungen und Mädchen!

Da die Klärung der Frage nach dem unterschiedlichen „Regionalen Intelligenz-Durchschnitt“ der Bundesländer eine Angelegenheit von großem öffentlichen Interesse ist, wäre es Sache des Bundesministeriums für Bildung und Forschung und ebenso Sache der Kultusministerien, vom Max-Planck-Institut für Bildungsforschung, dessen Arbeit von ihnen finanziert wird, die Erstellung und die Veröffentlichung solcher RID-Ranglisten einzufordern.

Fazit: *Die kognitiven Grundfähigkeiten* von Schülern als Ursache der *primären, begabungsbedingten Disparitäten* sind in den bisherigen Veröffentlichungen der Bildungsforschung nicht angemessen berücksichtigt worden. Erst mit ihrer Hilfe lassen sich auch die *tertiären, schulform-bedingten Disparitäten* genauer beschreiben.

6. Es fehlen zur Zeit genauere Informationen zur katastrophalen Situation der NRW-Gesamtschul-Oberstufen.

Vorbemerkung: Die katastrophale Situation der NRW-Gesamtschul-Oberstufen war, wie wir erfuhren, sowohl für das Fach Mathematik wie auch für das Fach Englisch schon im Jahre 1999 erkannt worden. Unseres Erachtens hätte damals schon, zeitgleich mit der Veröffentlichung des Mathematikvergleichs, eine deutliche Vorab-Information des NRW-Kultusministeriums stattfinden müssen, damit nicht ab 2007 noch acht weitere Jahrgänge von Gesamtschülern nach unzureichender Vorbereitung in die Probleme des NRW-Zentral-Abiturs geraten.

Ergebnisse des NRW-BIJU-Oberstufen-Vergleichs Mathematik von 1997

1997 sind vom MPIB im Rahmen des MPIB-Projektes „BIJU“ in Nordrhein-Westfalen an 12 Gesamtschulen und 19 Gymnasien im 12. und im 13. Jahrgang (zu gleicher Zeit) die Leistungen der Fächer Mathematik und Englisch verglichen worden (und in geringerem Umfang auch in Deutsch).

Zur Erinnerung: Wie von Baumert und Köller bereits in der Zeitschrift „Pädagogik“ (6/1998, S.17) mitgeteilt wurde, sind von ihnen im Rahmen des MPIB-Projektes „BIJU“ zwischen NRW-Gymnasiasten und gleich begabten NRW-Gesamtschülern am Ende des 10. Jahrgangs alarmierende Leistungsdifferenzen vorgefunden worden: *„Bei gleichen intellektuellen und sozialen Bedingungen beträgt der Leistungsvorsprung in Mathematik am Gymnasium mehr als zwei Schuljahre.“*

(Nach Ausweis der 1999 von Köller, Baumert und Schnabel veröffentlichten *„Tabelle 5“* zum NRW-Mittelstufen-Vergleich (s.u.,1999, S.404) waren die Leistungsunterschiede in Englisch mit 38,0 zu 37,6 Punkten übrigens ebenso groß wie in Mathematik.)

In Jahre 1999 veröffentlichten Köller, Baumert und Schnabel in der „Zeitschrift für Erziehungswissenschaft“ (Heft 3, S.385-422) einen Aufsatz über den im Jahre 1997 durchgeführten NRW-BIJU-Oberstufen-Vergleich Mathematik unter dem Titel *„Wege zur Hochschulreife: Offenheit des Systems und Sicherung vergleichbarer Standards“*. In diesem Aufsatz werden neben anderem auch die Ergebnisse eines Vergleichs von Gesamtschülern vorgestellt, die sowohl im 10. Jahrgang wie auch im 12. Jahrgang an der BIJU-Studie teilgenommen hatten.

Ergebnis: *„Festzuhalten ist, dass sich die Leistungsunterschiede zwischen Gesamtschulen und Gymnasien nicht nachweislich zwischen den Jahrgangsstufen 10 und 12 verändern. [...] Die Unterschiede bleiben in einer Größenordnung von gut einer SD.“* (Köller/Baumert/Schnabel 1999, S.408; SD=Standardabweichung)

An anderer Stelle wird mitgeteilt, *„dass man davon auszugehen hat, dass die Leistungsunterschiede von der 12. zur 13. Jahrgangsstufe konstant bleiben.“* (Köller/Baumert/ Schnabel 1999, S.406)

„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 Leistungsunterscheide erkennbar.“ (Köller, Baumert u.a. 2004, S.691)

Dieser Abstand von *„einer Standardabweichung“* bleibt also bis zum Ende des 13. Jahrgangs nahezu unverändert bestehen. Er entspricht dem Abstand von zwei Notenstufen der an Gymnasien üblichen Zensierung (Köller/Baumert/Schnabel 1999, S.410, *Abbildung 1*).

Das bedeutet: Was an NRW-Gesamtschulen *„befriedigend“* genannt wird, würde an NRW-Gymnasien mit einem *„mangelhaft“* bewertet. *„Lehrkräfte wählen offensichtlich für ihre Leistungsbewertung einen schulinternen Referenzrahmen.“* (Köller/Baumert/Schnabel 1999, S.409)

„Bemerkenswerterweise sind die Unterschiede in den kognitiven Grundfähigkeiten mit etwa einer halben Standardabweichung kleiner.“ (Köller/Baumert/Schnabel 1999, S.405). An NRW-Gymnasien hätten diese Gesamtschüler also höhere Leistungen und gerechtere Noten erreicht.

Olaf Köller beschreibt die Ergebnisse des NRW-BIJU-Oberstufen-Vergleichs Mathematik im „MPIB-Bildungsbericht 2003“ unter der Überschrift „Forschungsergebnisse aus der Gesamtschule“ (S.483) wie folgt: ***„In der gymnasialen Oberstufe finden sich in Mathematik Leistungsunterschiede zwischen beiden Schulformen, die in derselben Größenordnung liegen wie am Ende der Sekundarstufe I, das heißt, die Differenzen, die bereits beim Eintritt in die Oberstufe bestehen, wachsen nicht weiter an, reduzieren sich aber auch nicht.“***

Das ist im Grunde ein sehr deutlicher Hinweis auf die Qualität des Mathematik-Unterrichtes in den Oberstufen von NRW-Gesamtschulen. Eigentlich hätte infolge des sogenannten *„Matthäus-Effektes“* mit einer weiteren Öffnung der Leistungsschere gerechnet werden müssen. Die blieb aus. Daher ist die Ursache für die Leistungsunterschiede nicht in der Oberstufe von Gesamtschulen, sondern im niedrigen Fördereffekt ihrer 5. bis 10. Jahrgänge zu suchen. (Mt 25,29: *„Wer hat, dem wird noch hinzugegeben. Wer aber wenig hat, dem wird auch das, was er hat, noch genommen werden.“*)

Ergebnisse des NRW-BIJU-Oberstufen-Vergleichs Englisch von 1997

Im Jahre 2004 veröffentlichten Olaf Köller, Jürgen Baumert, Kai S. Cortina, Ulrich Trautwein und Rainer Watermann einen Aufsatz über den im Jahre 1997 durchgeführten NRW-BIJU-Oberstufen-Vergleich Englisch unter dem Titel: *„Öffnung von Bildungswegen in der Sekundarstufe II und die Wahrung von Standards“*, und zwar in der *„Zeitschrift für Pädagogik“* (Heft 5, S.679-700). Er enthält neben anderem folgende wichtige und höchst beunruhigende Ergebnisse:

An Gesamtschulen ereignet sich im Fache Englisch in der Oberstufe bis zum Ende des 13. Jahrgangs ein Kompetenz-Schwund von mehr als 40%: Offenbar verlernen Gesamtschüler in der Oberstufe, was sie in der Mittelstufe gelernt haben (und das war ja leider ohnehin schon nicht sehr viel).

Dieser Befund, der bisher noch nicht mit der erforderlichen Aufmerksamkeit registriert wurde, ist der *„Tabelle 5“* (S.695) des von Köller, Baumert u.a. 2004 veröffentlichten Berichtes zu entnehmen.

Die sprachliche Kompetenz im Fach Englisch war mit dem an amerikanischen Universitäten üblichen *„Test of English as a Foreign Language* (TOEFL)*“* überprüft worden. Der Mindestwert sind 500 TOEFL-Punkte. Mit ihm werden ausländische Studenten an einigen amerikanischen Universitäten bereits zum Studium zugelassen. Die meisten amerikanischen Universitäten verlangen jedoch mindestens 550 bis 600 Punkte.

Im zweiten Satz des ersten Abschnitts wird mitgeteilt, an Gymnasien seien im 12. Jahrgang die Leistungen der Schüler *„erwartungsgemäß etwas niedriger“*. Das ist eine Selbstverständlichkeit und daher unseres Erachtens wohl ein *„versteckter Hinweis“*. Denn es hätte anschließend dann ja auch noch mitgeteilt werden müssen, dass an NRW-Gesamtschulen - anders als an NRW-Gymnasien - die Englischleistungen im 12. Jahrgang *„erwartungswidrig deutlich höher“* sind als im 13. Jahrgang!

Nur bei dieser Betrachtungsweise wird erkennbar, dass der zweite Satz des nächsten Abschnitts die vorgefundenen, alarmierenden Leistungsdefizite in einer verharmlosenden Unvollständigkeit referiert.

Dort steht: *„Zudem lässt sich keine Leistungsverbesserung von der 12. zur 13. Jahrgangsstufe nachweisen.“* Das entspricht zwar der Wahrheit. Aber es ist weniger als die „halbe“ Wahrheit: Die „ganze“ Wahrheit lautet: **Von der 12. zur 13. Jahrgangsstufe ereignet sich an den NRW-Gesamtschulen im Fach Englisch eine *„Leistungsverschlechterung“* um mehr als 40%.**

Wegen dieser irreführenden sprachlichen Verschleierung von katastrophalen Leistungsdefiziten wird nachfolgend die Tabelle 5 mit dem Kommentar in Kopie wiedergegeben. (Hinweis: Es sind in diesem Aufsatz auch Untersuchungsergebnisse aus Baden-Württemberg veröffentlicht worden.)

Tab. 5: **Prozentualer Anteil von Schülerinnen und Schülern der 12. und 13. Jahrgangsstufe, die kritische Werte im TOEFL erreichen oder überschreiten, nach Kursniveau und Schulform in Nordrhein-Westfalen**

		Aufnahmekriterien amerikanischer Universitäten	
		≥ 500	≥ 550
12. Jahrgang	GK	52,6%	14,3%
Allgemein bildendes Gymnasium	LK	79,7%	37,1%
	Abgewählt	48,8%	10,7%
13. Jahrgang	GK	68,0%	28,8%
Allgemein bildendes Gymnasium	LK	87,4%	52,6%
	Abgewählt	44,0%	8,5%
12. Jahrgang	GK	6,7%	0,8%
Integrierte Gesamtschule	LK	27,4% !	8,9% !
	Abgewählt	10,9%	0,0%
13. Jahrgang	GK	6,5%	0,0%
Integrierte Gesamtschule	LK	15,8% !	2,6% !
	Abgewählt	5,7%	0,9%

LK: Leistungskurs; GK: Grundkurs

Sichtbar ist, dass Schülerinnen und Schüler der 13. Jahrgangsstufe an allgemein bildenden Gymnasien in Nordrhein-Westfalen vergleichbare Leistungen wie in Baden-Württemberg erreichen(vgl. Tabelle 3). In 12. Jahrgangsstufen allgemein bildender Gymnasien liegen die Leistungen erwartungsgemäß etwas niedriger.

Ganz anders fällt das Bild für Oberstufen an Gesamtschulen aus. Der Anteil der Untersuchungsteilnehmer, die kritische Werte erreichen, liegt deutlich unter den Vergleichswerten an beruflichen Gymnasien in Baden-Württemberg. Zudem lässt sich keine Leistungsverbesserung von der 12. zur 13. Jahrgangsstufe nachweisen.

Eine Auswertung der in *„Tabelle 5“* (S.695) dargestellten Leistungen hat dieses Ergebniss:

In den Leistungskursen von NRW-Gesamtschulen erreichten oder überschritten im 12. Jahrgang nach Ausweis der *„Tabelle 5“* lediglich 27,4% der Schüler den TOEFL-Wert 500, im 13. Jahrgang waren es dann nur noch 15,8% (**Leistungsabfall: -42,3%**).

Lediglich 8,9% von diesen Gesamtschülern erreichten oder überschritten im 12. Jahrgang auch den empfehlenswerteren TOEFL-Wert 550, im 13. Jahrgang waren es dann nur noch 2,6%. (**Leistungsabfall: -70,7%**)

In den Leistungskursen von NRW-Gymnasien erreichten oder überschritten im 12. Jahrgang nach Ausweis der *„Tabelle 5“* bei deutlich besserer Ausgangslage 79,7% der Schüler den kritischen TOEFL-Wert 500. Im 13. Jahrgang waren es 87,4% (**Leistungszuwachs: +9,6%**).

Von diesen Gymnasiasten erreichten bzw. überschritten im 12. Jahrgang 37,1% den empfehlenswerteren TOEFL-Wert 550. Im 13. Jahrgang waren es dann 52,5%. (**Leistungszuwachs: +41,5%**)

Es wurden in der Auswertung zunächst nur die Leistungskurse berücksichtigt, weil genaue Aussagen zur Leistungsentwicklung in Grundkursen wegen der Abwahlmöglichkeit nicht möglich sind.

Trotzdem ist das Grundkursergebnis eigentlich noch viel aufschlussreicher:

In den Grundkursen des 13. Jahrgangs von Gymnasien erreichen 28,8% der Schüler den empfehlenswerteren TOEFL-Wert 550, im 12. Jahrgang waren es erst 14,3%. (**Leistungszuwachs: +101,4%**)

In den Grundkursen des 12. Jahrgangs von Gesamtschulen erreichen 0,8% den TOEFL-Wert 550.

In den Grundkursen des 13. Jahrgangs von Gesamtschulen schaffte es kein einziger Schüler, **Leistungs-Abfall: total.**

„Ungefähr 90% der Abgänger aus integrierten Gesamtschulen bleiben unter dem kritischen Wert von 500." (Köller, Baumert u.a. 2004, S.695) Nach Ausweis der *„Tabelle 5"* (S.695) sind es bei den Gymnasiasten vermutlich weniger als 20%. Genauere Zahlen fehlen. Die entgegengesetzten Zahlangaben wären erheblich eindrucksvoller gewesen: nur 10% der Gesamtschüler erreichten oder überschritten im 13. Jahrgang den kritischen Mindestwert von 500 TOEFL-Punkten, bei den Gymnasiasten waren es vermutlich mehr als 80%.

Den höheren Wert von 550 TOEFL-Punkten erreichten oder überschritten im 13. Jahrgang lediglich 2,5% der Gesamtschüler. Bei den Gymnasiasten waren es 52,5% der Leistungskursschüler und dann noch einmal 28,8% der Grundkursschüler.

Bezüglich der unterschiedlichen Begabung der Oberstufenschüler von NRW-Gesamtschulen und NRW-Gymnasien gab es im Oberstufen-Vergleich Mathematik diese bereits erwähnte Auskunft:

„Bemerkenswerterweise sind die Unterschiede in den kognitiven Grundfähigkeiten mit etwa einer halben Standardabweichung kleiner." (Köller/Baumert/Schnabel 1999, S.405). Der Unterschied der Leistungen lag im Mittel *„bei ungefähr einer Standardabweichung"*. Das heißt: An NRW-Gymnasien hätten die Gesamtschüler auch im Fach Englisch deutlich bessere Leistungen erreicht.

Im Gegensatz zum NRW-Mittelstufenvergleich (vgl. Pädagogik 6/1998, S.17) sind aus dem NRW-Oberstufenvergleich Ergebnisse von *„Vergleichen vergleichbarer Schüler"* nicht veröffentlicht worden, obwohl auch für diese NRW-Oberstufenschüler Daten zu ihrem *„sozialen Hintergrund"* und zu ihrer *„kognitiven Grundfähigkeit"* vorlagen (S.694/695). Die Benachteiligung der NRW-Gesamtschüler hätte dann noch deutlicher nachgewiesen werden können.

Zur Erinnerung: Die Leistungsunterschiede zwischen NRW-Gymnasiasten und gleich begabten NRW-Gesamtschülern lagen bereits am Ende des 10. Jahrgangs mit einer Standardabweichung bei *„mehr als zwei Schuljahren"* (Baumert/Köller 1998, S.17; nach Ausweis einer Grafik waren es sogar mehr als drei Schuljahre!).

Vom MPIB wird die Tatsache, dass NRW-Gesamtschüler infolge des unzureichenden Mittelstufen-Trainings derart schlechte Werte erreichen, folgendermaßen kommentiert:

„Gerade die längsschnittlichen Befunde in NRW zeigen, dass die Öffnung alternativer Wege zur Hochschulreife langfristig wohl nur durchhaltbar ist, wenn in der Sekundarstufe I Maßnahmen ergriffen werden, um das Auseinanderklaffen der Leistungen zwischen den Schulformen zu reduzieren. Dies gilt insbesondere für Gesamtschulen in NRW. Allein der Versuch, Defizite zu Beginn der Oberstufe zu kompensieren, erscheint aufgrund der sehr großen Leistungsunterschiede nicht hinreichend zu sein. Die Befunde legen aber auch Überlegungen nahe, Korrekturmöglichkeiten zum Studienbeginn anzubieten." (Köller, Baumert u.a., 2004, S.698; nachträglich unterstrichen)

Gesamtschüler sollen also an der Universität nachholen, was ihre Altersgenossen bereits während ihrer Schulzeit an den Gymnasien lernen konnten und was dort das selbstverständliche *„Angebot des Hauses“* war?

Die von Köller und Baumert im Oerter/Montada 2002 ausgesprochene Empfehlung *„Frühe Differenzierung fördert leistungsstarke Schüler“*, die doch eine bereits mit dem 5. Jahrgang einsetzende schulform-bezogene Differenzierung als die effektivere Lösung dieser Probleme erscheinen lässt, wird nicht aufgegriffen.

Fazit: Schüler von NRW-Gesamtschulen werden in der Mittelstufe und in der Oberstufe, wie vom MPIB durch die BIJU-Oberstufen-Vergleiche nachgewiesen worden ist, nicht ihren Möglichkeiten entsprechend gefördert. Sie haben darum in der Berufs-Ausbildung und beim Studium, insbesondere bei Auslands-Studien, deutlich schlechtere Chancen als gleich begabte Altersgenossen, die NRW-Gymnasien besuchen.

Die NRW-Gesamtschulen genügen nicht mehr dem 1969 ausgerufenen Versprechen einer *„größeren Chancengleichheit“* durch eine begabungsgerechtere, *„bessere Förderung“*.

Die Ergebnisse der in den Jahren 2000/2001 durchgeführten postalischen Befragung von etwa 3.000 BIJU-Teilnehmern über deren weiteren Lebensweg sind vom MPIB noch nicht veröffentlicht worden. Die hier zu erwartenden Auskünfte würden diese Feststellung auf ihre Weise bestätigen.

„Integrierte Gesamtschulen sind schon ohne die Ganztagsbetreuung um 20% teurer als Gymnasien, bei bis zu 40% weniger Lernleistung.“ (Feststellung des Verfassers bei Gelegenheit eines Vortrags in Linz, Österreich, am 24.09.1999) NRW-Gesamtschulen sind inzwischen also, wie vom MPIB mit der Veröffentlichung der NRW-BIJU-Oberstufen-Vergleiche bestätigt wurde, auch ein Fall für den Finanzminister von NRW.

Vorbemerkung zum Folgenden: Im Teil C des KMK-Bildungsberichtes 2003 sollten *„zentrale Befunde zur Wirkung von Bildung und Ausbildung referiert werden“* (S.12). In diesem Zusammenhang wird in C2.4 (S.200/201) auch über die 1999 veröffentlichten Ergebnisse des NRW-BIJU-Oberstufen-Vergleichs Mathematik berichtet. Die katastrophalen Leistungsdefizite der NRW-Gesamtschul-Oberstufen werden dabei durch die Form der Darstellung irreführend verschleiert.

Im KMK-Bildungsbericht 2003 wird unter Verweis auf den von Köller, Baumert und Schnabel im Jahre 1999 veröffentlichten Bericht bezüglich der in den Oberstufen von Gesamtschulen und Gymnasien vorgefundenen Mathematik-Leistungen folgendermaßen referiert:

„<u>Hinsichtlich der Leistungsgewinne in der gymnasialen Oberstufe</u> ist die Befundlage sehr schmal, allein die BIJU-Studie lässt eine grobe Abschätzung der Zuwächse in den ersten beiden Jahren der gymnasialen Oberstufe zu. Köller, Baumert und Schnabel berichten für gymnasiale Oberstufen in Nordrhein-Westfalen getrennt für Schulformen (<u>Gymnasien vs. Gesamtschulen</u>) und Kursniveaus (Grund- vs. Leistungskurs) die Leistungsverläufe für Mathematik. Die Gewinne pro Schuljahr liegen dort über einer Standardabweichung, was für eine Beschleunigung der Zuwächse in der Oberstufe spricht. <u>Weiterhin zeigen sich bei den Zuwächsen keine systematischen Unterschiede zwischen Schulformen</u>, wohl aber zwischen Kursen zu Gunsten der Leistungskurse.“ (Bildungsbericht 2003, S.200/201; Unterstreichungen nachträglich eingefügt)

Dass es *„zwischen den Schulformen bei den Zuwächsen keine systematischen Unterschiede“* gebe, ist eine höchst missverständliche Mitteilung, obgleich sie im Wortlaut der Wahrheit entspricht.
^1

Die *„Unterschiede zwischen den Schulformen“* bleiben zwar während der Oberstufenzeit *„konstant“*, aber sie bleiben konstant auf der Höhe des bereits am Ende des 10. Jahrgangs vorliegenden Leistungs-Unterschieds von *„mehr als zwei Schuljahren“* bzw. *„gut einer Standardabweichung“* und entsprechen, wie bereits vorgetragen, zwei Notenstufen der an den Oberstufen von Gymnasien üblichen Zensierung.

Zwischenbilanz: Wenn im KMK-Bildungsbericht 2003 also mitgeteilt wird, es gäbe *„bei den Zuwächsen“* zwischen Oberstufen von NRW-Gymnasien und NRW-Gesamtschulen *„keine systematischen Unterschiede“*, dann entspricht das durchaus der Wahrheit. Aber es ist nicht die ganze Wahrheit, sondern - in diesem Falle - noch weniger als die halbe Wahrheit. Dennoch erweckt es den Eindruck, die ganze Wahrheit zu sein.

Nach klassischen Moral-Theorien handelt es sich hier um eine **„restrictio mentalis“**, um *eine im Geiste vollzogene Einengung, Beschränkung oder Zurückhaltung dessen, was eigentlich mitgeteilt werden müsste.* Der Begriff ist abgeleitet von lateinisch *„restringere: zurückbinden - beschränken - zurückhalten“* (Georges 1880 II, S.2107)

In der Vorbemerkung des Bildungsberichtes 2003 (Innenseite des vorderen Buchdeckels) steht zu lesen: *„Dieser Bericht wurde im Auftrag der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland erstellt. Die in dem Bericht dargestellten Ergebnisse und Interpretationen liegen in der alleinigen Verantwortung der Verfasserinnen und Verfasser.“* Damit sich derartiges nicht wiederholt, sollte also herausgefunden werden, wer diese verkürzten Darstellungen in den Text geschleust hat.

In Wahrheit ist es nämlich so:

Bereits am Ende des 10. Jahrgangs bestehen zwischen NRW-Gymnasiasten und *„vergleichbaren“* NRW-Gesamtschülern zum Beispiel in Mathematik schon sehr große Leistungsunterschiede: *„Bei gleichen intellektuellen und sozialen Bedingungen beträgt der Leistungsvorsprung in Mathematik am Gymnasium mehr als zwei Schuljahre.“* (Baumert/ Köller 1998, S.17)

Über Schüler, die 1995 auch im 10. Jahrgang schon an den BIJU-Untersuchungen teilgenommen hatten, wurde 1999 mitgeteilt: *„Festzuhalten ist, dass sich die Leistungsunterschiede zwischen Gesamtschulen und Gymnasien nicht nachweislich zwischen den Jahrgangsstufen 10 und 12 verändern. (Interaktion „Schulform x Jahrgangsstufe“: F [1, 448] = 1.1, ns). Die Unterschiede bleiben in einer Größenordnung von gut einer SD.“* (Köller/Baumert/Schnabel 1999, S.408; SD=Standardabweichung) - Man hat weiterhin davon auszugehen, *„dass die Leistungsunterschiede von der 12. zur 13. Jahrgangsstufe konstant bleiben.“* (Köller/Baumert/Schnabel 1999 S.406)

Es stimmt also: Das Lerntempo der Oberstufenschüler von NRW-Gesamtschulen und NRW-Gymnasien ist ohne *„systematische Unterschiede“*, aber es ist ein ähnliches Lerntempo auf sehr unterschiedlichen Leistungsebenen. Das heißt: In der Realität bestehen schon vor Beginn des 11. Jahrgangs erhebliche Leistungsunterschiede, die während der Oberstufe bei gleichem Lerntempo lediglich eine Parallelverschiebung erfahren. *„Im Mittel liegen die Differenzen bei ungefähr einer Standardabweichung.“* (Köller/Baumert/Schnabel 1999, S.405) Der Unterschied entspricht, wie erwähnt, dem Abstand von zwei Notenstufen der in den Oberstufen von Gymnasien üblichen Zensierung: Was an der Gesamtschule *„befriedigend“* genannt wird, würde an Gymnasien als *„mangelhaft“* benotet. (Köller/Baumert/Schnabel 1999, S.410, *„Abbildung 1: Zusammenhang (Regressionsgeraden) zwischen Noten und Testleistungen in Mathematik nach Schulform und Kursniveau.“*)

Fazit: Den Autoren des KMK-Bildungsberichtes 2003 ist die auf den Seiten 200/201 veröffentlichte Desinformation offenbar nicht aufgefallen und den Autoren des Bildungsberichtes 2006 offenbar auch nicht. Denn sonst wären diese Darstellungen, welche desaströse Missstände des Schulwesens von NRW verschleiern, im Bildungsbericht 2006 doch wohl korrigiert worden.

7. Es fehlt eine öffentliche Korrektur der im Teil C des KMK-Bildungsbericht 2003 enthaltenen, folgenreichen Unvollständigkeiten. Sie betreffen die Vorteile einer mit dem 5. Jahrgang einsetzenden frühen Differenzierung nach Fähigkeiten.

Vorbemerkung: Die 5. und 6. Jahrgänge sind in den letzten 35 Jahren in allen bundesländerübergreifenden Leistungs-Vergleichen ausgespart geblieben, obgleich sie doch die Weichenstellungen jeder Schullaufbahn enthalten.

Bei dieser Vorbemerkung berufen wir uns auf eine bereits erwähnte, unseren Vorwurf bestätigende Feststellung von Peter M. Roeder: *„Hinsichtlich der Schulleistungseffekte einer Verlängerung der Grundschuldauer liegt meines Wissens nur eine repräsentative Studie vor, nämlich eine Sekundäranalyse der um 1970 durch das Max-Planck-Institut für Bildungsforschung erhobenen Daten zur Leistungsentwicklung von Gymnasiasten des 7. Schuljahres (Roeder & Sang, 1991).“* Diese Feststellung hat Roeder in der durch Weinert und Helmke herausgegebenen Aufsatzsammlung *„Entwicklung im Grundschulalter“* (1997, S.407) veröffentlicht. Sie gilt auch heute noch.

Auf diesen von Roeder und Sang im Jahre 1991 veröffentlichten Aufsatz *„Über die institutionelle Verarbeitung von Leistungsunterschieden“* sowie auf die in ihm veröffentlichte *„Tabelle 1“* berufen sich Köller und Baumert im Oerter/Montada 2002 in ihrem Beitrag zur *„Entwicklung schulischer Leistungen“*. Dort (S.771) referieren sie unter der Überschrift *„Frühe Differenzierung fördert leistungsstarke Schüler“* folgende wichtige und offenbar immer noch gültige Erkenntnis aus dem MPIB-Projekt *„Schulleistung“* (1968-1970): ***„Zusammenfassend zeigt sich zumindest für das deutsche Schulsystem, dass bezogen auf die Fachleistungsentwicklung leistungsstarke Schüler von der Differenzierung im Sekundarbereich profitieren.“***

Zwischenbilanz: Die Wiedergabe dieser Erkenntnis des MPIB im KMK- Bildungsbericht 2003 ist unvollständig und daher irreführend missverständlich. Sie entspricht nicht dem Auftrag der Kultusministerkonferenz, *„durch eine tiefgreifende und auf empirische Daten sich stützende Analyse über Entwicklungen im Bildungswesen der Bundesrepublik Deutschland zu informieren“* (Vorwort der hessischen Kultusministerin Karin Wolff zum KMK- Bildungsbericht 2003, S.1).

Im 4. Abschnitt von Kapitel C2 (C2.4) geht es um die Beschreibung von *„Leistungszuwächsen“*. Auf der Seite 200 werden dort zunächst Befunde aus der TIMSS II 1997 vorgetragen. Dann folgt die Frage, *„ob die günstigeren Entwicklungsverläufe auf dem Gymnasium bloße Folge der Zusammensetzung leistungsstarker Schülerinnen und Schülern sind “*, und deren Beantwortung. Es wird dazu auf eine von Köller und Baumert im Jahre 2001 veröffentlichte Auswertung von BIJU-Daten verwiesen. Aber die nun folgenden Sätze stammen nicht aus dieser im Jahre 2001 veröffentlichten Auswertung von BIJU-Daten. Sie sind vielmehr wörtlich übernommen aus jenem Beitrag von Köller und Baumert im Oerter/Montada 2002 zur *„Entwicklung schulischer Leistungen“* (S.770/771), der auf Daten des MPIB-Projektes *„Schulleistung“* beruht und in Zusammenfassung auch Ergebnisse der Studie von 2001 referiert.

Im Oerter/Montada 2002 wird von Köller und Baumert in ihrem Beitrag über die *„Entwicklung schulischer Leistungen“* mitgeteilt, dass in anderen Ländern, zum Beispiel in Hongkong, eine *„Leistungsgruppierung nach Fähigkeiten“*, wie sie im deutschen gegliederten Schulwesen üblich ist, keinerlei Vorteile hat. Dann folgt der überraschende Hinweis, zumindest für Deutschland gelte das nicht. Denn hier profitieren leistungsstärkere Schüler *„unübersehbar“* von einer mit dem 5. Jahrgang einsetzenden frühen Differenzierung in getrennten Schulformen.

Bewiesen wird das anhand der Daten zum Leistungsstand in Deutsch, Englisch und Mathematik, die im Rahmen des MPIB-Projektes „Schulleistung“ am Anfang des 7. Jahrgangs bei Gymnasiasten aus Bundesländern mit vierjähriger Grundschule und bei Berliner Schülern nach sechsjähriger Grundschule vorgefunden worden waren.

Im KMK-Bildungsberichtes 2003 werden in jenem Kapitel C2.4, in dem über *„Leistungszuwächse“* berichtet werden sollte, zwar größere Teile der Ausführungen dieses Beitrags von Köller und Baumert aus dem Oerter/Montada 2002 im vollen Wortlaut zitiert. Aber es fehlen eigenartigerweise alle Informationen, die in jenen Kreisen, die eine sechsjährige oder gar eine zehnjährige Grundschule propagieren, höchst *„unerwünscht“* wären: Es fehlen alle wichtigen Passagen, in denen Köller und Baumert die Vorteile des mit dem 5. Jahrgang einsetzenden gegliederten Schulwesens und den niedrigen Fördereffekt sechsjähriger deutscher Grundschulen beschreiben. Die Darstellung erweckt vielmehr den Eindruck, als *„ob die günstigeren Entwicklungsverläufe auf dem Gymnasium“* gerade nicht die Folge einer *„Leistungsgruppierung nach Fähigkeiten“* wären. - Und die Zitate sind nicht als Zitate gekennzeichnet worden!

Offensichtlich sind den neun Hauptautoren und den drei Koautoren des KMK-Bildungsberichtes 2003 diese sich als folgenreiche Informations-Lücken auswirkenden Auslassungen nicht aufgefallen.

Die Unvollständigkeit der Darstellung dieser hoch relevanten Erkenntnisse der Bildungsforschung im KMK-Bildungsbericht 2003 hatte heikle Folgen: Im Jahre 2005 ist in Mecklenburg-Vorpommern die schulart-unabhängige Orientierungsstufe eingeführt worden - und in Schleswig-Holstein wäre im Jahre 2005 beinahe sogar die zehnjährige Grundschule eingeführt worden.

Der Vorgang des selektiven Ausblendens wird nachfolgend in synoptischer Gegenüberstellung dokumentiert. Die weißen Felder „markieren“ jene Passagen, die ausgelassen worden sind. Die Synopse enthält auch den ersten Teil dieses Abschnittes aus dem Oerter/Montada. Er steht unter der Überschrift *„Frühe Differenzierung fördert leistungsstarke Schüler“*. Diese Einführung und die aufschlussreiche Tabelle 23.2. sind im KMK-Bildungsbericht 2003 trotz ihres hohen Informationsgehaltes nicht abgedruckt worden. Zitiert wird dort lediglich aus dem Exkurs, zwar Wort für Wort, aber das nur unvollständig. Dem Verfasser dieser „folgenreichen Unvollständigkeiten“ hat, wie die Synopse zeigt, der vollständige Text aus dem Oerter/Montada vorgelegen. Der außerordentlich wichtige letzte Satz des Exkurses fehlt im Bildungsbericht 2003 völlig.

Oerter/Montada 2002	**KMK-Bildungsbericht 2003**
Oerter/Montada: *„Entwicklungspsychologie“* 2002, Kapitel 23: Olaf Köller und Jürgen Baumert, S.756ff: *„Auswirkungen der Leistungsgruppierung“* (S.770/771)	Hermann Avenarius, Hartmut Ditton u.a.: „Bildungsbericht für Deutschland - Erste Befunde“ Leske + Budrich, Opladen 2003 (Federführung: Deutsches Institut für Internationale Pädagogische Forschung (DIPF)) **Kapitel:** C2.4 ***„Leistungszuwächse“*** (S.200/1)
„Frühe Differenzierung fördert leistungsstarke Schüler. Implizieren diese internationalen Befunde, dass die Fähigkeitsgruppierungen wenigstens im Hinblick auf optimale Leistungsförderungen ineffektiv sind, so ergibt sich in deutschen Studien ein abweichendes Bild. Roeder (1997; vgl. auch Roeder & Sang 1991) reanalysierte Daten einer vom Max-Planck-Institut für Bildungsforschung Ende der 60er Jahre durchgeführten Gymnasiastenstudie mit einer für alle Länder der alten Bundesrepublik repräsentativen Stichprobe. Ausgangspunkt der Arbeit war die Fragestellung, 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. Jahrgangstufe 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	$\bar{x}$	38.8	38.0	49.5	1049
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	

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

Unübersehbar zeigen sich deutliche Leistungsvorteile der Gymnasiasten aus den Bundesländern mit vierjähriger Grundschule. Die frühe Differenzierung scheint also leistungsstärkere Schüler in allen drei Fächern besser zu fördern.

Exkurs: *„Effekte der Leistungsgruppierung in Deutschland"*

Eine methodisch elaboriertere Studie zu Effekten der Leistungsgruppierung im deutschen Schulsystem stammt von Köller und Baumert (2001).

Die Autoren wählten ebenso wie in den oben beschrieben Arbeiten von Marsh et al. (2000) und Yang et al. (1999) einen mehrebenenanalytischen Auswertungsansatz.

Basierend auf einer Stichprobe von N=2730 Schülerinnen und Schülern aus 107 Schulen wurden die Leistungsverläufe in Mathematik untersucht und auf die individuelle Ausgangsfähigkeit, die auf Schulebene aggregierte Ausgangsfähigkeit und die Schulform zurückgeführt. Die Analysen zeigten das folgende Ergebnismuster. Bei Kontrolle der übrigen Prädikatoren hatte die individuelle Ausgangsfähigkeit in Klasse 7 einen deutlich positiven Effekt auf die Leistung in Klasse 10 (β=.50). Die Schulform hatte ebenfalls einen substantiellen Einfluss:

Bei Kontrolle der individuellen und der auf Schulebene aggregierten Ausgangsleistung lag die in Klasse 10 erreichte Leistung am Gymnasium um mehr als eine halbe Standardabweichung über der Realschule und fast eine Standardabweichung über der Hauptschule. Bei Kontrolle von individueller Ausgangsfähigkeit und Schulform zeigte sich nur ein schwacher positiver Effekt (β=.10, *ns*) der auf Schulebene aggregierten Ausgangsleistung in der 7. Jahrgangsstufe, d.h. innerhalb einer Schulform gab es nur unbedeutende Unterschiede in der Leistungsentwicklung zwischen leistungsstärkeren und -schwächeren Schulen. Dieses Ergebnis ist keinesfalls trivial, sofern man bedenkt, dass die Leistungsvarianz zwischen den Schulen innerhalb von Schulformen manchmal größer ist als die Varianz zwischen den Schulformen (z.B. Fend, 1982).

Köller und Baumert (2001) interpretieren ihre Ergebnisse dahingehend, dass es offenbar stärker die besondere Instruktionskultur am Gymnasium und weniger die Leistungsgruppierung *per se* sei, die sich leistungsfördernd auswirken könnte.

Diese besondere Instruktionskultur ist dabei möglicherweise weniger Folge eines makroadaptiven Verhaltens der Lehrkräfte auf die vorgefundene Leistungsstärke in den jeweiligen Gymnasialklasse, als vielmehr Folge der schulformspezifischen Lehrerausbildung im deutschen Lehrerbildungssystem.

Zusammenfassend zeigt sich zumindest für das deutsche Schulsystem, dass bezogen auf die Fachleistungsentwicklung leistungsstarke Schüler von der Differenzierung im Sekundarbereich profitieren.

(Kopie u. Abschrift; die letzten Zeilen mit Absicht in Fettdruck)

Köller und Baumert haben mit den BIJU-Daten für das Fach Mathematik systematisch untersucht, ob die günstigeren Entwicklungsverläufe auf dem Gymnasium bloße Folge der Zusammensetzung leistungsstarker Schülerinnen und Schüler sind.

Basierend auf einer großen Stichprobe wurden die Leistungsverläufe in Mathematik von der 7. bis zur 10. Jahrgangsstufe untersucht und auf die individuelle Ausgangsfähigkeit, die auf Schulebene aggregierte Ausgangsfähigkeit und die Schulform zurückgeführt.

Bei Kontrolle von individueller Ausgangsfähigkeit und Schulform zeigte sich kein Effekt der auf Schulebene aggregierten Ausgangsleistung in der 7. Jahrgangsstufe, d.h. innerhalb einer Schulform gab es nur unbedeutende Unterschiede in der Leistungsentwicklung zwischen leistungsstärkeren und -schwächeren Schulen. Dieses Ergebnis ist keinesfalls trivial, sofern man bedenkt, dass die Leistungsvarianz zwischen den Schulen innerhalb von Schulformen manchmal größer ist als die Varianz zwischen den Schulformen.

Köller und Baumert interpretieren ihre Ergebnisse dahingehend, dass es offenbar stärker die besondere Instruktionskultur am Gymnasium und weniger die Leistungsgruppierung *per se* sei, die sich leistungsfördernd auswirken könnte.

(Abschrift)

(Das von Köller und Baumert erwähnte *„Bundesland mit sechsjähriger Grundschule"* ist Berlin)

Der Beitrag von Köller und Baumert im Oerter/Montada 2002 zu den Vorteilen einer frühen Differenzierung war jenen Personen, von denen diese Darstellung im KMK-Bildungsbericht 2003 stammt, offensichtlich bekannt. Dennoch erweckt die Art der Wiedergabe in ihrer Unvollständigkeit auch hier den Eindruck, als ob *„die Leistungsgruppierung nach Fähigkeiten“* keinen Einfluss auf die Leistungsentwicklung hätte.

Zwischenbilanz: Die wichtigsten Informationen aus der Darstellung von Köller und Baumert zu den positiven *„Auswirkungen der Leistungsgruppierung“* (Oerter/Montada 2002, S.766-771) und zu den Vorteilen einer frühen, mit dem 5. Jahrgang einsetzenden Differenzierung sind im KMK-Bildungsbericht 2003 offensichtlich mit Vorsatz ausgelassen worden, obgleich sie in dem Streit um die sechsjährige Grundschule endlich Klarheit gebracht hätten.

„Dieser Bericht wurde im Auftrag der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland erstellt. Die in dem Bericht dargestellten Ergebnisse und Interpretationen liegen in der alleinigen Verantwortung der Verfasserinnen und Verfasser.“ (Vorbemerkung zum KMK-Bildungsbericht 2003, auf der Innenseite des vorderen Buchdeckels) Es könnte also auch hier wieder herausgefunden werden, auf welche Weise diese folgenreichen Unvollständigkeiten in den Text geraten sind.

Die im „Teil C“ des KMK-Bildungsberichtes 2003 enthaltenen irreführend unvollständigen Informationen sind in einem Bericht von Autoren, bei denen eigentlich ein hoher Informationsstand und eine solche Vorgänge verhindernde Kompetenz erwartet werden dürften, absolut erwartungswidrig.

Die Annahme, derartig verkürzte Informationen seien mit Absicht zugelassen worden, darf aus Respekt vor den Autoren gar nicht erst erwogen werden. Also bleibt als einzige Erklärung für diese Unzulänglichkeiten und Unvollständigkeiten im Teil C des KMK-Bildungsberichtes 2003 die Annahme einer *„unverschuldeten Ahnungslosigkeit“*, verursacht durch die *zurückhaltende* Informations-Politik des Max-Planck-Institutes für Bildungsforschung.

Fazit: Schon die Tatsache, dass gleich an zwei Stellen des KMK-Bildungsberichtes 2003 ein dreistes Ausblenden wichtiger Befunde riskiert werden konnte, ohne dass dies auffiel, ist ein hinreichender Grund für die Feststellung: Sowohl die Kultusministerien wie auch die Fachkollegen sind über lange Jahre vom Max-Planck-Institut für Bildungsforschung nicht gründlich genug informiert worden.

Unseres Erachtens gehört es zu den wichtigsten Aufgaben eines Institutes für Bildungsforschung, seine Erkenntnisse zeitnah, vollständig, eindeutig und in allgemeinverständlicher Form an die Kultusministerien, an alle Fachkollegen und an die Öffentlichkeit weiterzureichen, damit derartige Informationsdefizite und Fehlentscheidungen gar nicht erst zustande kommen.

Das Max-Planck-Institut für Bildungsforschung ist 1963 gegründet worden wegen des Versprechens, für Theorie und Praxis des Unterrichtens *„eine empirisch abgesicherte Basis“* zu schaffen.

Daran wollten wir mit Nachdruck erinnert haben.

Unser Vorwurf lautet: Am Max-Planck-Institut für Bildungsforschung wurden wichtige Ergebnisse aus aufwändigen MPIB-Studien nicht oder nur zum Teil und oft erst mit großer Verspätung bekannt gemacht - aber dann keineswegs mit einer dem Aufwand und der Bedeutung entsprechenden, breitenwirksamen Allgemeinverständlichkeit. Stattdessen wurden sie als *„versteckte Botschaften“* in der Fachliteratur deponiert, wo sie unbeachtet und ohne Wirkung blieben.

In Dänemark würde die Zurückhaltung von Forschungsergebnissen sofort zum Gegenstand von Untersuchungen werden. Bei diesem Hinweis beziehen wir uns auf die im Januar 1998 von der Deutschen Forschungsgemeinschaft (DFG) aus gegebener Veranlassung veröffentlichten ***„Empfehlungen zur Sicherung guter wissenschaftlicher Praxis“*** und die dort (S.33) zitierten dänischen Regelungen.

Derartige Regelungen fehlen für Deutschland. In den Empfehlungen der Deutschen Forschungsgemeinschaft (DFG) von 1998 werden (S. 33) die dänischen Regelungen zitiert. Nach diesen Regelungen würden *„Fälle selektiven Ausblendens und Verschweigens ‚unerwünschter' Ergebnisse"* in Dänemark *„eine förmliche Untersuchung grundsätzlich rechtfertigen oder erfordern"*.

Was im deutschen Schulwesen seit 1976 geschehen ist, lässt sich vielleicht so veranschaulichen:

Einmal angenommen, das *„Max-Planck-Institut für medizinische Forschung"* in Heidelberg hätte herausgefunden, dass eine traditionelle Behandlungsmethode der Kinderheilkunde im hiesigen Klima erheblich wirksamer ist als eine von den zuständigen Ministerien auf Anraten eben dieses Institutes mit großem Aufwand eingeführte neue Behandlungsmethode:

Wie würde die Öffentlichkeit wohl reagieren, wenn sie erführe, dass dieses Institut über lange Jahre - aus welchen Gründen auch immer - nicht im Stande gewesen wäre, den zuständigen Ministerien und den Eltern der mit jener neuen Methode behandelten Kinder die Korrektur seiner Empfehlungen auf wirksame Weise bekannt zu machen? Hatte die Öffentlichkeit nicht das Recht, zu erwarten, dass dieses Institut seine problematischen Befunde umgehend veröffentlicht und vor der weiteren Anwendung der von ihm zuvor empfohlenen, angeblich besseren Behandlungsmethode gewarnt hätte?

Dann hätten die zuständigen Ministerien und Behörden unter Wahrnehmung ihrer Fürsorgepflicht die weitere Anwendung der neuen, aber weit weniger wirksamen Behandlungsmethode umgehend untersagen können.

Zusammenfassung:

Der von der Kultusministerkonferenz in Auftrag gegebene KMK-Bildungsbericht 2003 enthält in sehr wichtigen Angelegenheiten Informationen, die offensichtlich mit Vorsatz unvollständig waren. Das Vertrauen der Autoren dieses Berichtes ist durch das Einschleusen dieser folgenreich-unvollständigen Informationen ausgenutzt worden.

Die unvollständigen Informationen konnten aber nur darum in den KMK-Bildungsbericht 2003 geraten, weil die Mitarbeiter des Deutschen Institutes für Internationale Pädagogische Forschung, das auch für diesen Bildungsbericht schon die Federführung hatte, offenbar seit langen Jahren vom Max-Planck-Institut für Bildungsforschung nicht gründlich genug über die Ergebnisse von MPIB-Studien informiert worden sind. Das Ansehen des Deutschen Institutes für Internationale Pädagogische Forschung (DIPF) ist durch diese *zurückhaltende* Informations-Politik des Max-Planck-Institutes für Bildungsforschung beschädigt worden.

Auch die Kultusministerien sind vom MPIB seit 30 Jahren nicht gründlich genug über die Ergebnisse des MPIB-Projektes *„Schulleistung"* (1968-1970) und des MPIB-Projekts „BIJU" (1991-2002) informiert worden. Denn sonst wäre es doch gewiss in Mecklenburg-Vorpommern nicht zur Einführung der Regionalschule und der schulart-unabhängigen Orientierungsstufe gekommen. Durch die *zurückhaltende* Informations-Politik des Max-Planck-Institutes für Bildungsforschung ist also auch das Ansehen der Kultusministerien beschädigt worden.

Die Verlässlichkeit der deutschen Bildungsforschung wurde unseres Wissens zum letzten Mal vor mehr als 25 Jahren überprüft. Das war, als Professor Dr. Heinrich Wottawa (Bochum) auf Anregung von Professor Dr. Aloys Regenbrecht (Münster) die Ergebnisse der wissenschaftlichen Begleitung des Gesamtschulversuchs nachgerechnet hat. Ergebnis: Wottawas Aufsatz *„Die Leistungsfähigkeit der NRW-Gesamtschulen - eine kritische Betrachtung der empirischen Ergebnisse"* (1980) und sein Aufsatz über *„Die Kunst der manipulativen Berichtlegung in der Evaluationsforschung"* (1981).

In diesem zweiten Aufsatz werden von Wottawa verschiedene Beispiele aufgelistet, wie durch entsprechende Formulierungen *„unerwünschte"* Befunde verschleiert werden können. Gegenstand seiner Untersuchungen war der Abschlussbericht der *„Wissenschaftlichen Beratungsgruppe Gesamtschulversuch NW"*. Und eben dieser Bericht machte es möglich, dass in Nordrhein-Westfalen im Jahr 1982 die Gesamtschule neben den bestehenden Schulen als Regelschule eingeführt werden konnte.

Recklinghausen, Oktober 2006