

15. Juni 2004

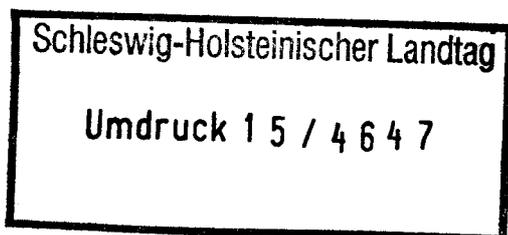
Sehr geehrter Herr Schmidt

Auf bei dir von Ihnen erbeten Stellungnahme.  
Wenn sie nicht der vorgeschriebenen Form genügt,  
dann lassen Sie mich das wissen.

Vorerst bin ich telefonisch um über mein  
Handy erreichbar. Aber ich bin Mitte Juli  
für einige Tage wieder in Recklinghausen,  
und könnte dann entsprechend reagieren.

Mit den besten Wünschen für Ihre Arbeit,

U. Sprenger



## „Frühe Differenzierung fördert leistungsstarke Schüler“

Ein Beitrag des Max-Planck-Institutes für Bildungsforschung (MPIB) zur Schulform-Debatte

Die neue Ausgabe des pädagogischen Standardwerkes „Entwicklungspsychologie“ von Oerter und Montada (Beltz Verlage 2002) enthält - mit eben dieser Feststellung - einen Beitrag der Professoren Dr. Olaf Köller (bis 2002 Mitarbeiter am MPIB) und Dr. Jürgen Baumert (seit 1996 Direktor am MPIB) zum Thema „**Entwicklung schulischer Leistungen**“.

**Köller und Baumert referieren zunächst (S.769), dass in anderen Ländern ein gegliedertes Schulsystem mit seiner Gruppierung der Schüler nach Fähigkeiten keine Vorteile bringt. Dann aber folgt die überraschende Mitteilung (S. 770), das sei zumindest in Deutschland anders: „Implizieren diese internationalen Befunde, dass die Fähigkeitsgruppierungen wenigstens im Hinblick auf optimale Leistungsförderungen ineffektiv sind, so ergibt sich in deutschen Studien ein abweichendes Bild.“**

Belegt wird die Feststellung mit Daten aus der ersten großen Studie des 1963 gegründeten Max-Planck-Institutes für Bildungsforschung (MPIB). Es handelt sich um das MPIB-Projekt „Schulleistung“, auch „**Gymnasiastenstudie**“ genannt. Für diese sehr aufwendige Studie sind vom MPIB 1968 bis 1970 in den zehn alten Ländern der Bundesrepublik und in West-Berlin etwa 14.000 Gymnasiasten von etwa 450 Gymnasien je drei Tage zu Beginn und am Ende des 7. Jahrgangs untersucht worden, und zwar im Hinblick auf ihre Leistungen in den Fächern Deutsch, Englisch, Mathematik, außerdem auch noch im Hinblick auf ihre intellektuellen Grundfähigkeiten, auf ihr Freizeitverhalten sowie im Hinblick auf den Bildungsstand der Eltern und der Großeltern-Generation. Letzteres, um bei der Auswertung der Daten auch den Einfluss des „familiären Hintergrundes“ einschätzen zu können.

Eine Teil-Auswertung dieser Daten war von Professor Dr. Peter M. Roeder (1973 bis 1995 als Vorgänger Baumerts Direktor am MPIB) und seinem Mitarbeiter Dr. Fritz Sang zunächst 1991 und dann 1997 noch einmal in gekürzter Form von Roeder veröffentlicht worden. Auf diese Veröffentlichungen berufen sich Köller und Baumert bei ihrem Beweisgang (S. 770): „*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 Schulleistung hat. Dazu wurden die Leistungen von Gymnasiasten in der 7. Jahrgangsstufe in Deutsch, Mathematik und Englisch aus einem Bundesland mit sechsjähriger Grundschule mit denen der übrigen Bundesländer mit vierjähriger Grundschule verglichen. Die Tabelle 23.2 zeigt die Resultate der Studie.*“ (Das Land mit sechsjähriger Grundschule war West-Berlin.)

Die Tabelle 23.2 wird im Text nicht weiter kommentiert. Daher ist nur für jemanden, der mit dieser Art Statistiken einigermaßen vertraut ist, folgendes zu erkennen: **In Englisch und Mathematik liegt der Leistungsrückstand nach sechsjähriger Grundschule bei etwa einer Standardabweichung. Das ist - wie aus anderen MPIB-Studien erschlossen werden kann - ein Leistungsrückstand von etwa anderthalb Schuljahren!**

Die Veröffentlichung von 1991 enthält auch Angaben über die weitere Entwicklung der beiden Schülergruppen und über den am Ende des 7. Jahrgangs vorgefundenen Leistungsstand: „*Im Englischen liegt der Unterschied immer noch bei etwa einer Standardabweichung, bleibt also fast unverändert bestehen. Im Mathematik- und Deutschunterricht geht die Differenz auf etwa eine Drittel Standardabweichung zurück, ist also immer noch auch praktisch bedeutsam. (...) Die Aussagen der Lehrer über ihre Unterrichtsplanung für das 7. Schuljahr und über die Themen und Aufgaben, die sie späteren Schuljahren zuweisen, deuten darauf hin, dass der Leistungsausgleich zwischen beiden Gruppen langsamer vonstatten geht wird, als man bei gleich befähigten Schülern vermuten würde.*“ (S.167/68)

Vor diesem Hintergrund bekommt die folgende, von Köller und Baumert vorgelegte Zusammenfassung von Ergebnissen des Projektes „Schulleistung“ ein erheblich höheres Gewicht: **„Unübersehbar zeigen sich deutliche Leistungsvorteile der Gymnasiasten aus den Bundesländern mit vierjähriger Grundschule. Die frühere Differenzierung scheint also leistungsstärkere Schüler in allen drei Fächern besser zu fördern.“** (S. 770)

Offenbar sind die mit Hilfe der Gymnasiasten-Studie gewonnenen Ergebnisse nach wie vor gültig. Denn die beiden Autoren halten es nicht für erforderlich, sie mit aktuelleren Daten zum niedrigeren Fördereffekt von undifferenzierten Lerngruppen zu bestätigen, etwa mit den entsprechenden Daten des 1991 vom MPIB gestarteten Projektes BIJU („Bildungsprozesse und psychosoziale Entwicklung im Jugendalter und jungen Erwachsenenalter“), anhand der dort 1991 im 7. Jahrgang in Berlin und NRW vorgefundenen Leistungsunterschiede in den Fächern Deutsch, Mathematik und Englisch. Wohl aber berichten sie (S. 778) unter Bezug auf neuere MPIB-Studien, dass auch **„die Intelligenzentwicklung am Gymnasium deutlich günstiger“** ist: **„Mit dem günstigeren Lernklima am Gymnasium gehen offenbar nicht nur bessere Schulleistungsentwicklungen einher, sondern auch eine besondere Ausschöpfung kognitiver Begabungsreserven“.**

Die von Köller und Baumert vorgelegten, aus MPIB-Studien gewonnenen Erkenntnisse zum höheren Fördereffekt des gegliederten Schulwesens gipfeln (S. 771) in der Feststellung:

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

#### **Diese Feststellung erfordert unseres Erachtens folgende Konsequenzen:**

- 1. Es sollte niemand mehr von der Einführung sechsjähriger Grundschulen eine Verbesserung der Lernleistungen deutscher Schüler erwarten - und Entsprechendes ausrufen. Wer aber weiterhin daran festhält, ist entweder schlecht informiert oder er nimmt die Befunde des MPIB nicht ernst.**
- 2. Eine über den 4. Jahrgang hinausgehende Unterrichtung leistungsstarker Schüler in undifferenzierten, leistungsgemischten Lerngruppen ist nach Bekanntwerden dieser MPIB-Befunde als eine vorsätzliche Benachteiligung leistungsstarker Schüler zu betrachten - und daher einzustellen.**
- 3. Auch Realschüler sind leistungsstärkere Schüler. Auch sie profitieren von einer frühen Differenzierung und von einer schulformbezogenen Ausbildung ihrer Lehrer (Oerter/Montada 2002 S. 771). Das stellt die in Regionalschulen, Regelschulen, Mittelschulen und Haupt-Realschulen übliche, erst nach dem 6. Jahrgang einsetzende Differenzierung in Frage, ebenso auch die Tatsache, dass es in NRW und anderen Ländern eine schulform-bezogene Ausbildung der Realschullehrer nicht mehr gibt.**
- 4. Nicht nur die leistungsstärkeren, sondern auch die leistungsschwächeren Schüler profitieren von einer frühen Differenzierung. Köller und Baumert verweisen (S. 772) auf die Arbeiten von Schwarzer und Jerusalem. Dort wird nachgewiesen, dass leistungsschwächere Schüler in leistungs-gemischten Lerngruppen hohen psychischen Strapazen ausgesetzt sind. In den leistungs-homogeneren Gruppen von Hauptschulen bleiben ihnen die täglichen Unzulänglichkeiterfahrungen erspart. Sie können hier ein höheres Selbstwertgefühl entwickeln. Diese Chance darf ihnen durch die Einführung einer Einheitsschule nicht genommen werden.**

Ulrich Sprenger, Arbeitskreis Gesamtschule e. V., Juni 2004

**Der Arbeitskreis Gesamtschule e. V.** (Sitz in Recklinghausen) existiert seit 1994. Seine Mitglieder sind in der Mehrzahl Lehrerinnen und Lehrer, die an Gesamtschulen unterrichten oder dort unterrichtet haben. Aufgrund ihrer praktischen Erfahrungen und gestützt auf Ergebnisse der Bildungsforschung halten sie die integrierte Gesamtschule in Deutschland unter den hier gegebenen Bedingungen nach wie vor für **„ein nicht zu haltendes Versprechen“** und für **„eine pädagogische Fehlkonstruktion“**. Sie hat sich in den 30 Jahren ihres Bestehens trotz aller Begünstigungen und Nachbesserungen nicht als **„die pädagogisch und volkswirtschaftlich effektivere Organisationsform von Schule“** erwiesen, als die sie 1969 vom Deutschen Bildungsrat empfohlen worden ist. (Weitere Informationen unter: [www.ak-gesamtschule.de](http://www.ak-gesamtschule.de))

# AKG | intern

Mitteilungen des Arbeitskreises Gesamtschule e.V.

Juni 2004

Sehr geehrte Damen und Herren  
Liebe Mitglieder

Wir sind alle hereingelegt worden: Die, die sich seit langen Jahren für die Einführung der Gesamtschule engagiert haben, ebenso wie die, die sie zu verhindern suchten. Die polarisierenden Schulformdebatten, „die Grabenkämpfe der 70er Jahre“ und viele Fehlinvestitionen hätten uns erspart werden können. Denn bereits 1969 lagen am Max-Planck-Institut für Bildungsforschung (MPIB) Daten vor, an Hand deren spätestens 1975 hätte gezeigt werden können, dass leistungsstarke Schüler in den leistungs-heterogenen Lerngruppen der 5. und 6. Jahrgänge von Gesamtschulen und Orientierungsstufen nicht begabungsgerecht gefördert werden. An Hand eben dieser Daten von 1969 und an Hand der BIJU-Daten von 1991/95 kommen die Professoren Dr. Olaf Köller (bis 2002 Mitarbeiter am MPIB) und Dr. Jürgen Baumert (seit 1996 Direktor am MPIB) im Oerter/Montada des Jahres 2002 - unter der Überschrift **„Frühe Differenzierung fördert leistungsstarke Schüler“** - zu folgenden Feststellungen: *„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.“* (S.770) - *„Mit dem günstigeren Lernklima am Gymnasium geben offenbar nicht nur bessere Schulleistungsentwicklungen einher, sondern auch eine besondere Ausschöpfung kognitiver Begabungsreserven.“* (S.778) - *„Zusammenfassend zeigt sich zumindest für das deutsche Schulsystem, dass - bezogen auf die Fachleistungsentwicklung - leistungsstarke Schüler von der Differenzierung im Sekundarbereich profitieren.“* (S.771) Das sind sehr wichtige Beiträge des MPIB zur Schulform-Debatte.

Spätestens bis 1974 hätten, wie uns auf Anfrage wiederholt bestätigt wurde, die Daten jenes auch „Gymnasiasten-Studie“ genannten MPIB-Projektes „Schulleistung“ ausgewertet und veröffentlicht werden können, noch vor dem blamablen Scheitern der Koop-Schule (1978) und vor der Anerkennung der Gesamtschule als „gleichwertiger“ Schulform (KMK 1982). Die Entwicklung des deutschen Schulwesens hätte einen günstigeren Verlauf genommen!

**„Frühe Differenzierung fördert leistungsstarke Schüler.“** - Nach Ausweis von KMK-Statistiken wurden im Jahre 2000 etwa 40 % aller deutschen Schülerinnen und Schüler auch in den 5. und 6. Jahrgängen noch in leistungs-gemischten Lerngruppen unterrichtet. Die vom MPIB veröffentlichten Befunde berechtigen zu der Schlussfolgerung: **Der Unterricht in leistungs-gemischten Lerngruppen ist eine der Ursachen für „das miserable Abschneiden“ der deutschen Schülerschaft bei den internationalen Leistungsvergleichen der PISA-Studie.**

Im Oerter/Montada 2002, (S. 769/770) verweisen Köller und Baumert zwar darauf, dass in anderen Ländern eine *„Gliederung des Schulwesens nach Fähigkeitsgruppen“* keine Vorteile bringt. Dies gelte aber nach Ausweis jener MPIB-Studie von 1969 nicht für Deutschland. Sie erhärten

diese Mitteilung durch Ergebnisse des MPIB-Projektes „Schulleistung“ und des MPIB-Projektes „BIJU“. Wer sich also auch jetzt noch für die Einführung der sechsjährigen Grundschule oder der Einheitsschule engagiert und dazu auf die PISA-Ergebnisse von Finnland und Schweden verweist, zeigt nur, dass er schlecht informiert ist - oder die Befunde des Max-Planck-Institutes für Bildungsforschung nicht ernst nimmt.

Über einen erweiterten Verteiler gingen die letzten Ausgaben des „AKG-intern“, in denen diese Dinge ausführlich dargestellt waren, an über 100 Adressaten aus Politik, Wissenschaft und Medien - und an alle Kultusminister, mit mancher Resonanz. Der bemerkenswerteste Kommentar stammt von einer der angeschriebenen Personen aus dem Bereich der Wissenschaft:

*„Die Tatsache, dass Ergebnisse, die 1969 über Schulleistungsvergleiche erhoben wurden, unter Becker, dem damaligen MPIB-Direktor, nicht veröffentlicht worden sind, hat zu tun mit den Empfehlungen des Deutschen Bildungsrates zu 40 Gesamtschulversuchen! Die von Becker vorbereitete und durchgesetzte Empfehlung wäre nämlich konterkariert worden, hätte man die Ergebnisse des Leistungsvergleichs gleichzeitig bekannt gemacht!“*  
(„gleichzeitig“ meint: gleichzeitig mit den 1969 gestarteten Gesamtschulversuchen.)

Echo und Resonanz, aber nicht die überfällige Entrüstung der Öffentlichkeit! So bleiben wegen der andauernden Vorherrschaft wohlgemeinter pädagogischer Wunschträume in Deutschland seit 30 Jahren Schüler und Lehrer auch weiterhin unter dem Niveau ihrer Möglichkeiten. Seit 30 Jahren gibt es in Deutschland keine vorurteilsfreie, realitätsnahe Diskussion über den unterschiedlichen Fördereffekt der verschiedenen Schulformen, weil Informationen der Bildungsforschung ausbleiben.

**So aber durfte und so darf mit der Zukunft von Kindern, mit den Hoffnungen der Eltern und mit der Einsatzbereitschaft von Lehrern nicht umgegangen werden!**

***Die nachfolgenden Ausführungen behandeln Sie bitte vertraulich, sie sind nicht zur Veröffentlichung bestimmt:*** Beratungen im Vorstand des Arbeitskreises Gesamtschule e. V. und in seiner Mitgliederversammlung ergaben: Es ist sehr schwer, sich einen Überblick zu verschaffen über die Unvollständigkeit der verschiedenen MPIB-Veröffentlichungen aus dem Forschungsbereich „Schule und Unterricht“. Denn es fehlt eine zusammenfassende, allgemeinverständliche kritische Darstellung der vier wichtigsten Studien mit ihren Ergebnissen, aber auch mit ihren möglichen, jedoch zurückgehaltenen Ergebnissen. Dies erklärt wohl das Ausbleiben der längst fälligen öffentlichen Entrüstung. Wir wollen uns daher daran machen, eine solche Übersicht zu erstellen, wenn möglich bis Ende März 2005. - Titel:

## **Der unkontrollierte Verfall des deutschen Bildungswesens**

**Erster Versuch einer Dokumentation zur Forschungspolitik und zur Informationspolitik des Max-Planck-Institutes für Bildungsforschung, die Leistungsentwicklung an weiterführenden Schulen betreffend.**

1. Kapitel: Vor- und Frühgeschichte des MPIB
2. Kapitel: Das MPIB-Projekt „Schulleistung“ (1968 – 1970)
3. Kapitel: Das MPIB-Projekt „Hauptschule“ (1980)
4. Kapitel: Das MPIB-Projekt „BIJU“ (1991 gestartet)
5. Kapitel: Die Informations-Defizite aus PISA 2000

Es werden zunächst die einzelnen Projekte des MPIB und ihre Ziele beschrieben. Dann werden die bisher veröffentlichten Ergebnisse vorgestellt und anschließend die vielen wichtigen Informationen, die an Hand des vorliegenden Datenmaterials möglich gewesen wären, aber nicht veröffentlicht wurden. Wir hoffen durch diese Darstellung (mit entsprechenden

Grafiken) endlich eine öffentliche Reflexion über die folgenreiche Zurückhaltung bildungspolitisch hochbedeutsamer Informationen in Gang bringen zu können.

Zentraler Ansatz unserer Kritik: Obwohl aus dem MPIB-Projekt „Schulleistung“ und aus dem MPIB-Projekt „BIJU“ auch Daten zu den kognitiven Grundfähigkeiten der Schüler vorliegen, gibt es bis heute keine Untersuchungen über das Verhältnis von Begabung und Leistung bezüglich der am Anfang des 7. Jahrgangs vorgefunden Leistungen. Für die Fächer Deutsch, Englisch und Mathematik wäre das möglich. Auch für Deutsch liegen solche Daten vor (vgl. BIJU 94). Das ist nicht bekannt.

**Erst jene Leistungsunterschiede, die bei einem „Vergleich vergleichbarer Schüler“ am Anfang des 7. Jahrgangs vorgefunden wurden, geben genauere Auskunft über den unterschiedlichen Fördereffekt der 5. und 6. Jahrgänge von Schulformen und Schulsysteme.**

Solche Vergleiche von Schülern, welche „*ähnliche intellektuelle und soziale Ausgangsbedingungen*“ haben, sind erheblich präziser als jene 1991 vorgelegten pauschalen Vergleiche, die Köller und Baumert im Jahre 2002 für ihren Oerter/Montada-Beitrag benutzt haben!

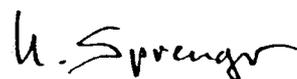
In der Zeitschrift „Pädagogik“ (6/98 S.17) wurde von Baumert und Köller mitgeteilt, dass NRW-Gymnasiasten gegenüber vergleichbaren NRW-Gesamtschülern, mit denen sie am Anfang des 7. Jahrgangs noch denselben Leistungsstand hatten, am Ende des 10. Jahrgangs z. B. „*in Mathematik einen Wissensvorsprung von mehr als zwei Jahren*“ hatten. Dies wird dort interpretiert als ein Beweis für den (von der GEW angezweifelten) geringeren Fördereffekt der Gesamtschule. Aber es wird nicht interpretiert als **ein Beweis für das Scheitern der Fachleistungsdifferenzierung**. Denn diese „*gymnasial befähigten*“ Gesamtschüler hatten ja am Anfang des siebten Jahrgangs noch den gleichen Leistungsstand wie die gleich begabten Gymnasiasten und sind ab da in den anspruchsvolleren Erweiterungskursen der Gesamtschulen unterrichtet worden. Wie in den „Empfehlungen des Deutschen Bildungsrates“ von 1969 versprochen, sollte aber die Gesamtschule gerade durch die Fachleistungsdifferenzierung „*volkswirtschaftlich und pädagogisch effektiver*“ sein als das gegliederte Schulwesen. Dass dies ein nicht zu haltendes Versprechen war, ist also seit 1998 zu erkennen.

Mit den BIJU-Daten wären auch Auskünfte darüber möglich, wie unterschiedlich sich Leistungen und Selbstwertgefühl von leistungsschwächeren Schülern an Gesamtschulen und an Hauptschulen entwickeln. Die geduldete Zurückhaltung dieser Auskünfte ist der eigentliche Skandal der Gesamtschul-Bewegung. - So viel als erste Andeutungen. - Eine MPIB-Studie über die Entwicklung des Selbstwertgefühls und der Leistungen von leistungsstarken Gymnasiasten an leistungsstarken Gymnasien liegt bereits vor!

Der nächste „**AKG-intern**“ erscheint im Oktober 2004. Thema: „**Die Kunst der manipulativen Berichterlegung in der Evaluationsforschung**“ - Zusammenfassung und Ergänzung eines Kommentars von Professor Heinrich Wottawa (Bochum) zum Bericht der Raschert-Kommission (Schriftenreihe des KM Nr. 38/1979) über den Gesamtschulversuch in NRW aus dem Jahre 1981 (Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 1/1981, S. 45 - 60). Außerdem der Abdruck der „**Vier Thesen zum Thema Gesamtschule**“, die vor zehn Jahren zur Gründung des Arbeitskreises führten

Auch im Namen des Vorstands  
die besten Wünsche für den Sommer  
und die kommende Urlaubszeit!

Freundliche Grüße



P. S.: Als Anlage die kurzgefasste Ausarbeitung eines der Themen, die ich mir für den Ruhestand aufgespart hatte. Für den multikulturellen Alltag an den Schulen ist sie vielleicht nicht uninteressant.

## Sehr geehrte Damen und Herren

Anbei unser Angebot: Wir unterstützen Sie bei Ihrer Arbeit durch die regelmäßige Zusendung unseres Informationsmaterials und durch die von uns vermittelten Kontakte. Sie Ihrerseits unterstützen uns durch einen Beitrag zu den nicht unerheblichen Kosten, die beim Zusammentragen dieser Informationen entstehen.

Sie haben also die Möglichkeit, in einer Art Abonnement stets genau so gut informiert zu sein wie die Mitglieder des Arbeitskreises. Zeitgleich und mit demselben Verteiler würde Ihnen der „AKG-intern“ zugehen, gegebenenfalls zusammen mit anderen aktuellen Informations-Materialien.

Der „AKG-intern“ erscheint mindestens viermal im Jahr.

Sollten Sie an der Zusendung unserer Informationsmaterialien interessiert sein, dann schicken Sie bitte den beiliegenden Antrag an die Geschäftsstelle oder an meine Adresse zurück. Die übrigen Anträge können Sie an Leute weitergeben, von denen Sie vermuten, dass sie gleichfalls an unseren Informationen interessiert sind.

Weitere und frühere Informationen erhalten Sie im Internet unter:

[www.ak-gesamtschule.de](http://www.ak-gesamtschule.de)

Dort finden Sie unter anderem unsere kommentierte Liste von gesamtschulkritischen Texten (Stichwort: „Literatur“, z.Zt. 161 Titel). Auf Abruf senden wir Ihnen - gegen Vorab-Erstattung der Kosten durch Briefmarken - Kopien von jenen Texten zu, die wir aus dieser Zusammenstellung weiterreichen können bzw. weiterreichen dürfen.

Wenn Sie Informationen haben, von denen Sie glauben, dass sie auch für andere interessant sind, dann lassen Sie die uns doch bitte zukommen, damit wir sie über den „AKG-intern“ weiterreichen können.

Freundliche Grüße,





Rolf Oerter  
Leo Montada  
(Hrsg.)

# ENTWICKLUNGSPSYCHOLOGIE

*5., vollständig überarbeitete Auflage (2002)*

## Kapitel 23 Entwicklung schulischer Leistungen

Olaf Köller & Jürgen Baumert

**BELTZPVU**

# Kapitel 23

## Entwicklung schulischer Leistungen

Olaf Köller & Jürgen Baumert

Die erfolgreiche lebenslange Entwicklung in einer sich rasant verändernden modernen Industrienation ist ohne Zweifel an breite intellektuelle Fähigkeiten und Fertigkeiten gebunden. Berufliche Anforderungen, aber auch der gesamte Freizeitbereich mit seinen immensen technologischen Veränderungen setzen auf Seiten aller Individuen einer modernen Gesellschaft ein Wissensrepertoire voraus, dessen Erwerb ohne institutionalisierte Bildungsprozesse in Schulen, Hochschulen und sonstigen Weiterbildungsstätten weitgehend unmöglich ist (vgl. Baumert et al., 2000a, 2000b; Geary, 1995, 1996; Helmke & Weinert, 1997). Der Schule als der vielleicht bedeutsamsten Bildungsinstitution kommt so die Rolle zu, Kindern und Jugendlichen systematische Lerngelegenheiten zu bieten, in denen sie in einem breiten Fächerkanon ein Wissensfundament aufbauen, das sie auf den Übergang in den Beruf und die Auseinandersetzung mit zunehmend komplexer werdenden Alltagsanforderungen vorbereitet.

### DEFINITION

Wir wollen im Folgenden unter **Schulleistungen** das auf den Lehrplänen basierende deklarative und prozedurale Wissen in verschiedenen Domänen verstehen, dessen Erwerb zu einem erheblichen Teil an Lerngelegenheiten im jeweiligen Fachunterricht gebunden ist.

Natürlich kommt der Schule nicht das Monopol bei der Wissensvermittlung zu. Es ist vielmehr das Zusammenspiel zwischen internalen Ressourcen des Kindes, den elterlichen Erziehungsbemühungen **und** schulischen Lernangeboten, das erfolg-

reiche Bildungsprozesse ermöglicht. Dabei ist die Bedeutung der verschiedenen Agenten (Schüler, Eltern, Schulen) für die Lernprozesse in unterschiedlichen Lebensabschnitten stärker bzw. schwächer. Grundlegende Prozesse des Spracherwerbs sind beispielsweise im Säuglings- und Kleinkindalter sehr stark an Reifungsprozesse des Kindes und seine Interaktionen mit den Eltern bzw. Kindergartenerzieherinnen gebunden (vgl. Schneider, 2001). In späteren Jahren übernimmt die Schule die Hauptfunktion beim Erwerb von Lesekompetenzen und der Schriftsprache. Schließlich verbessern Schulkinder mehr oder weniger selbstgesteuert ihre muttersprachlichen Kompetenzen durch Leseaktivitäten in der Freizeit, die allerdings erheblich durch das Angebot im elterlichen Haus beeinflusst sein können.

**Kernbereiche schulischer Bildung.** Mittlerweile haben sich vier Kernbereiche schulischer Bildung herauskristallisiert, in denen ein hinreichendes Maß an Kenntnissen notwendige Voraussetzung für eine gelingende Berufs- und Lebensperspektive ist. Es sind dies grundlegende Kompetenzen in der Muttersprache, der Mathematik, den Naturwissenschaften und Englisch bzw. der ersten Fremdsprache.<sup>1</sup> Ohne fundiertes Wissen in diesen Bereichen können im Extremfall nicht einmal die Einrichtungen des Wohlfahrtsstaates in Anspruch genommen werden (vgl. Klieme et al., 2000). Vor diesem Hintergrund der besonderen Relevanz schulisch erworbenen Wissens für eine erfolgrei-

<sup>1</sup> In der Regel ist die erste Fremdsprache Englisch, deren Bedeutung als *lingua franca* ihren Erwerb in der Schule unumgänglich macht.

che Entwicklung über die Lebensspanne sollen in diesem Kapitel aktuelle Arbeiten zur Entwicklung von Schulleistungen vorgestellt werden. Richtigerweise müsste man eigentlich von „Schülerleistungen“ sprechen, impliziert doch der Begriff „Schulleistung“, dass es sich um Leistungen der Schule handelt. Förderleistungen der Schule firmieren allerdings üblicherweise unter Namen wie „Schulleffizienz“ oder „Schuleffektivität“ (*school effectiveness*; vgl. Schnabel, 1998a; Scheerens & Bosker, 1997).

Zusätzlich wird ein Schwerpunkt auf die Rolle der Schule für erfolgreiche bzw. erfolglose Leistungsentwicklungen gelegt.

Die weiteren Ausführungen beschäftigen sich mit den folgenden Themen:

- (1) Theoretische bzw. konzeptionelle Probleme bei der Erforschung von Schulleistungen und ihrer Entwicklung im Laufe der Kindheit und des Jugendalters,
- (2) Entwicklungsverläufe von Schulleistungen in verschiedenen Fächern,
- (3) Die besondere Rolle der Leistungsdifferenzierung in der Sekundarstufe I für die individuelle Schulleistungsentwicklung,
- (4) Werden Schulleistungen bzw. ihre Entwicklung überhaupt maßgeblich durch schulische Lerngelegenheiten beeinflusst?
- (5) Das Zusammenspiel der Leistungsveränderung mit der Entwicklung lernfördernder individueller Variablen,
- (6) Methodische Probleme bei der Erforschung der Schulleistungsentwicklung.

## 1 Die fehlende Verankerung von Schulleistungen in psychologischen Theorien

Im Gegensatz zu prominenten Konstrukten wie der Intelligenz fehlt es an klaren psychologischen Theorien, die Schulleistungen hinsichtlich ihres Inhalts, ihrer Struktur, ihrer Genese und ihrer Entwicklung eindeutig beschreiben. Intelligenz-

theorien wie beispielsweise die Arbeit von Thurstone (1938) zu den *primary mental abilities* spezifizieren eine bestimmte Zahl von mehr oder weniger unabhängigen Dimensionen, die klar definiert werden und für deren Erfassung entsprechende Maße abgeleitet, entwickelt und validiert werden. Mit Hilfe dieser Instrumente (z.B. dem Prüfsystem für Schul- und Bildungsberatung PSB von Horn, 1969) ist es vergleichsweise einfach, die Entwicklung der Intelligenzdimensionen zu untersuchen und anschließend zu beschreiben. Die verwendeten Aufgaben in solchen Tests sind oftmals dieselben für verschiedene Altersgruppen bzw. variieren in ihrer Komplexität und Schwierigkeit über die Altersgruppen (wie im Falle des Kognitiven Fähigkeitstest von Heller & Perleth, 2000). Man verfällt hier üblicherweise nicht dem Glauben, dass sich die Konstrukteigenschaften über die Zeit verändern.

**Unter Schulleistung kann Verschiedenes verstanden werden.** Bei Schulleistungen liegt der Fall etwas anders, wie es beispielsweise Helmke und Schrader (1998, S. 60) ausführen: „Unter *Schulleistung* (SL) kann sehr verschiedenes verstanden werden, z.B. Leistungen unterschiedlicher Schüler und Schulklassen, prozedurales oder deklaratives Wissen, fachspezifisches Wissen oder überfachliche Fähigkeiten.“ Die Spezifizierung dessen, was genau unter Schulleistungen zu verstehen ist, erfolgt typischerweise nicht aus psychologischer, sondern aus didaktischer Perspektive. Konkret sind es die Lehrpläne (Curricula), die letztendlich auf einer inhaltlichen Ebene für jedes Fach und für jede Schulform in jedem Bundesland beschreiben, welche Inhalte die Schüler in welcher Jahrgangsstufe lernen sollen.<sup>2</sup>

**Die Anforderungen wachsen mit dem Alter der Schüler.** Mit jedem neuen Schuljahr und in jeder

<sup>2</sup> Psychologische Theoriebildungen zu Schulleistungen beschränken sich üblicherweise auf individuelle, gesellschaftliche und institutionelle Determinanten von Schulleistungen, wobei stillschweigend übergangen wird, worum es sich im Sinne psychologischer Konstrukte konkret bei Schulleistungen handelt.

Schulform verändern sich die curricularen Vorgaben dessen, was gelernt werden soll. Verlangen beispielsweise die Lehrpläne der Klassenstufe 5 für Mathematik, dass in der Geometrie einfache Körper (Rechteck und Quadrat) behandelt und von den Schülern beherrscht werden sollten, so geht es in der 10. Jahrgangsstufe im Geometrieunterricht um trigonometrische Funktionen (Sinus- und Kosinussatz mit Anwendung auf Dreiecksberechnungen). Gute Geometriekenntnisse in der 5. Jahrgangsstufe bedeuten dementsprechend, dass ein Schüler den Flächeninhalt eines Rechtecks berechnen kann, in der 10. Jahrgangsstufe, dass er aus einem gegebenen Winkel und einer gegebenen Seite ein Dreieck konstruieren kann.

**Spiegeln gute Fachleistungen in jedem Alter die gleiche latente Dimension wider?** Inwieweit beide Leistungen bzw. Kenntnisse noch Ausdruck der selben latenten Dimension sind, deren Veränderung sich über 5 Schuljahre empirisch untersuchen lässt, bleibt in bisherigen Forschungsarbeiten weitgehend offen. Wenn man allerdings die starke, empirisch überprüfbare Annahme macht, dass sich die Geometrieleistungen in der 5. und 10. Jahrgangsstufe auf einer gemeinsamen Dimension abbilden lassen (hierfür gibt es durchaus Evidenz, wie Analysen von Watermann, 2001 belegen), so wird deutlich, dass diese Dimension weniger ein psychologisches Merkmal im engeren Sinne ist, sondern vielmehr das erreichte Curriculum im Geometrieunterricht repräsentiert. Noch komplizierter wird das Ganze, wenn man sich vor Augen führt, dass die Geometrie ja nur ein Teilbereich der Mathematik ist und sich die Frage stellt, inwieweit Geometrieleistungen und Leistungen in anderen Teilbereichen der Mathematik (z.B. Algebra oder elementare Wahrscheinlichkeitsrechnung) auf eine gemeinsame latente Dimension zurückführbar sind.

**Eine einheitliche Theorie der Schulleistung kann es nicht geben.** Im Sinne psychologischer Theorienbildungen macht dies deutlich, dass es keine Theorie der Schulleistung bzw. der Schulleistungsentwicklung im engen Sinne geben kann.

Theoretisch kann der Gegenstand „Schulleistung“ am ehesten mit Bezug auf Lerntheorien bzw. die Lernpsychologie angegangen werden, indem man die Entwicklung von schulischen Leistungen ganz allgemein als Lernen im Sinne der relativ dauerhaften Veränderung von kognitiven Strukturen versteht (Schiefele, 1996; Skowronek, 1983). Für die empirische Untersuchung der Schulleistungsentwicklung resultieren aus diesen konzeptuellen Schwierigkeiten auch einige methodische Probleme, die sich auf die Validität und die Dimensionalität von Schulleistungstests beziehen. Auf beide Schwierigkeiten und weitere Probleme bei der Erfassung von Schulleistungen soll am Ende dieses Beitrags eingegangen werden.

## 2 Die Entwicklung von Schulleistungen in der Kindheit und im Jugendalter

### 2.1 Lerngelegenheiten und Leistungsentwicklung

Schulleistungen und damit auch ihre Entwicklung sind eng an Lerngelegenheiten gekoppelt, für die je nach Fach die Schule mehr oder weniger das Monopol hat. Dies hat für die Untersuchung von Schulleistungsverläufen über die Kindheit und Jugend wichtige Implikationen. Bieten Schulen oder andere außerschulische Agenten fortlaufend systematische Lerngelegenheiten, so ist mit einem permanenten Anstieg der Leistungen zu rechnen, sofern der Lernprozess kumulativ in dem Sinne ist, dass bereits früher erworbene Kenntnisse im Unterricht vertieft und erweitert und nicht etwa vergessen werden. Die Anknüpfungsfähigkeit neuer Inhalte an bereits erworbene Kenntnisse entscheidet darüber, ob das Fachwissen vertieft bzw. konsolidiert wird und ob die kognitive Repräsentation des neu erworbenen Wissens gelingt (vgl. hierzu z.B. Renkl, 1996). Handelt es sich um weitgehend kompartimentalisierte Unterrichtsinhalte, d.h. der neue Stoff knüpft inhaltlich kaum an früher Gelerntes an, so

können (a) auf Seiten des bereits erworbenen Wissens Vergessensprozesse einsetzen und (b) die neuen Inhalte schwieriger abgespeichert werden. Die aktuelle Forschung (im Überblick Renkl, 1996) hat sehr eindrucksvoll die Rolle des Vorwissens für gelingendes schulisches Lernen hervorgehoben, wobei der positive Einfluss umso stärker ist, je kumulativer ein Fach ist, d.h. je weniger kompartimentalisiert die verschiedenen Stoffgebiete sind.

**Die Naturwissenschaften kommen im Lehrplan zu kurz.** Die Abhängigkeit der schulischen Leistungsentwicklung von vorhandenen Lerngelegenheiten hat auch Implikationen für die Rolle der Wochenstundenzahl und Kontinuität, mit der Fächer unterrichtet werden. Während die Kernfächer Deutsch und Mathematik beginnend mit der ersten Grundschulklasse durchgehend bis zum Ende der Sekundarstufe I unterrichtet werden (mit wenigstens vier Wochenstunden in der Grundschule und wenigstens drei Wochenstunden in der Sekundarstufe I), werden die Naturwissenschaften und die Fremdsprachen nicht kontinuierlich über die gesamte Grund- und Mittelstufenzeit angeboten. Für die Naturwissenschaften gilt, dass sich der Sachkundeunterricht der Grundschule oftmals auf die Biologie beschränkt und die Physik und Chemie überhaupt nicht berücksichtigt sind. Nicht selten wird der Unterricht sogar durch heimatkundliche Inhalte dominiert (vgl. Einsiedler, im Druck). Ein Hauptgrund dafür ist, dass man orientiert an Piagets Theorie der kognitiven Entwicklung (s. den Beitrag von Montada in diesem Band) lange Zeit glaubte, dass Kindern im konkret-operatorischen Stadium der kognitiven Entwicklung die Voraussetzungen für den Erwerb der formal-abstrakten Unterrichtsinhalte der Physik und Chemie fehlen.<sup>3</sup> Basierend auf den Arbeiten Wygotskis argu-

mentiert Stern neuerdings (vgl. Stern, im Druck; Stern et al., im Druck), dass es bei gelingender Nutzung graphisch-visueller Repräsentationsformen im Sachkundeunterricht sehr wohl möglich sei, auch abstraktere Inhalte, wie sie Gegenstand der Physik sind, zu lehren.

**Unregelmäßiger Unterricht begünstigt Vergessensprozesse.** Im Laufe der Sekundarstufe I werden alle drei Naturwissenschaften nicht durchgehend mit größeren Stundenzahlen unterrichtet. Zum Teil findet integrierter Naturwissenschaftsunterricht statt, zum Teil werden die Fächer epochal unterrichtet, d.h. für einen gewissen Zeitraum wird nur Biologie unterrichtet, dann über ein längeres Intervall nur Physik usw. Der Chemieunterricht setzt sehr spät in der Sekundarstufe I ein, teilweise erst in der 9. Jahrgangsstufe, so dass dort ein systematischer Wissensaufbau in der Schule überhaupt nicht stattfindet. Der unregelmäßige Unterricht in einzelnen Fächern scheint in erheblichem Maß die Entwicklungsverläufe der Schulleistungen zu beeinflussen und zwar in dem Sinne, dass deutliche Vergessensprozesse einsetzen. Hierfür reichen offenbar relativ kurze Zeitperioden wie die Schulferien aus, in denen kein Unterricht stattfindet. Bryk und Raudenbush (1989) konnten diese Vergessensprozesse anhand von längsschnittlichen Daten für Mathematik und Muttersprache (*reading*) belegen. Im Sommerhalbjahr, in dem längere Ferienzeiten waren, waren die Lernkurven der Schüler in beiden Fächern abgeflacht, in Mathematik kam es teilweise sogar zu Verlusten. Die geringeren Verluste in Englisch lassen sich dadurch erklären, dass in der Muttersprache mehr außerschulische Lerngelegenheiten gegeben sind, gerade auch, wenn es um das Lesen geht, das bei Bryk und Raudenbush untersucht wurde.

**Frühzeitiger Beginn des Fremdsprachenunterrichts.** Der Fremdsprachenunterricht setzte traditionell im deutschen Schulsystem in der 5. Jahrgangsstufe ein, mittlerweile haben aber verschiedene Bundesländer den Englisch-Unterricht in der Grundschule eingeführt (beipielsweise in Bayern ab der 3. Jahrgangsstufe). Das spätere Ein-

<sup>3</sup> Als ein zweiter Grund für die fehlende Präsenz vor allem der Physik im Grundschulbereich kam und kommt hinzu, dass den entsprechenden Lehrern oftmals die fachlich-didaktische Qualifikation für beide Fächer fehlt.

setzen des Fremdsprachenunterrichts rührte aus der Annahme, dass die sichere Beherrschung der Muttersprache notwendige Voraussetzung für den gelingenden Fremdsprachenerwerb ist. Aufgrund aktueller Arbeiten zur Sprachentwicklung bei zweisprachig aufwachsenden Kindern (Brisk, 1997) ist man allerdings von diesen Annahmen abgerückt, was die Einführung des Englischunterrichts im Grundschulbereich forciert hat.

### Schulleistungen in der Grundschule

#### Erwartung fachspezifischer Leistungsverläufe.

Im Sinne der skizzierten Argumentation kann man also je nach Fach differentielle Entwicklungsverläufe der Leistungen erwarten, die entsprechend den unterschiedlichen Lerngelegenheiten in den Sprachen und Mathematik steiler, in den Naturwissenschaften flacher verlaufen sollten. Für die Grundschulzeit wissen wir relativ wenig über längerfristige Schulleistungsverläufe, da aktuelle Studien sich stärker mit der Prognosekraft von vorschulisch erhobenen Merkmalen (beispielsweise Intelligenz) für später gezeigte Schulleistungen interessieren. Die empirisch fruchtbarste deutsche Untersuchung im Grundschulbereich ist mit Sicherheit die SCHOLASTIK-Studie<sup>4</sup> (vgl. Weinert & Helmke, 1997).

**Ergebnisse der SCHOLASTIK-Studie.** Analysen der SCHOLASTIK-Daten von Schneider et al. (1997) ergaben für die Rechtschreibleistung den in Abbildung 23.1 skizzierten Entwicklungsverlauf. Die Kurve wurde auf der Basis mehrerebenenanalytischer Verfahren gewonnen (zu den Details s. Schneider et al., 1997, S. 123f). Der negativ beschleunigte und offensichtlich nicht lineare Kompetenzzuwachs über die Zeit war hoch signifikant und die Autoren resümieren diese Befunde wie folgt: „Die Befunde lassen insgesamt den Schluss zu, dass sich für die große Mehrheit der Grundschul Kinder von der zweiten bis zur vierten

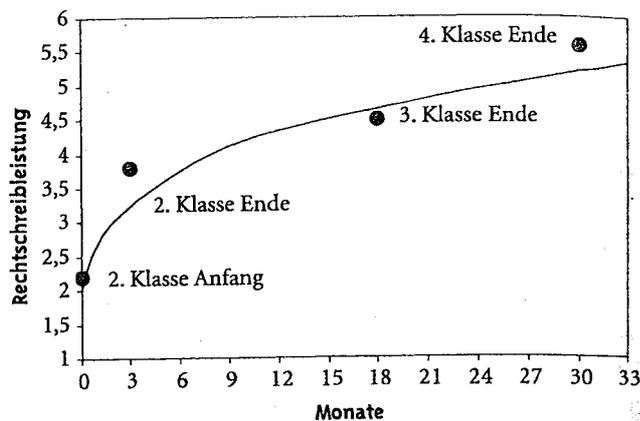


Abbildung 23.1. Entwicklung der Rechtschreibleistungen im Grundschulalter (aus Schneider et al., 1997, S. 124)

Klasse beträchtliche Fortschritte in den Rechtschreibkompetenzen registrieren lassen ...“ (Schneider et al., 1997, S. 128). Ähnlich große Fortschritte berichtet Reusser (1997) für den Bereich der Grundschulmathematik, konkret zur Veränderung der prozentualen Lösungshäufigkeiten von Textaufgaben während der Grundschulzeit (vgl. Tabelle V.2 in Reusser, 1997).

#### Schulleistung in der Sekundarstufe I

Zur Beschreibung der Entwicklungsverläufe von curricular verankerten Schulleistungen im Sekundarschulbereich I bietet sich die Kohorten-Längsschnittstudie „Bildungsprozesse und psychosoziale Entwicklung im Jugendalter und jungen Erwachsenenalter“ (BIJU) an, die seit 1991 am Max-Planck-Institut für Bildungsforschung in Berlin in Kooperation mit dem Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN) in Kiel durchgeführt wird. In Abbildung 23.2 sind die Leistungsverläufe für Englisch, Mathematik und Physik vom Beginn der 7. bis zum Ende der 10. Jahrgangsstufe, differenziert nach Schulformen, dokumentiert.<sup>5</sup> Die Datenerhebungen fanden zu vier Zeitpunkten statt. In allen Fächern

4 Das Akronym SCHOLASTIK steht für Schulorganisierte Lernangebote und Sozialisation von Talenten, Interessen und Kompetenzen.

5 Wachstumskurvenanalysen der Daten zeigten, dass sich über den hier untersuchten Zeitraum lineare Zuwachsmodele zufriedenstellend anpassen lassen.

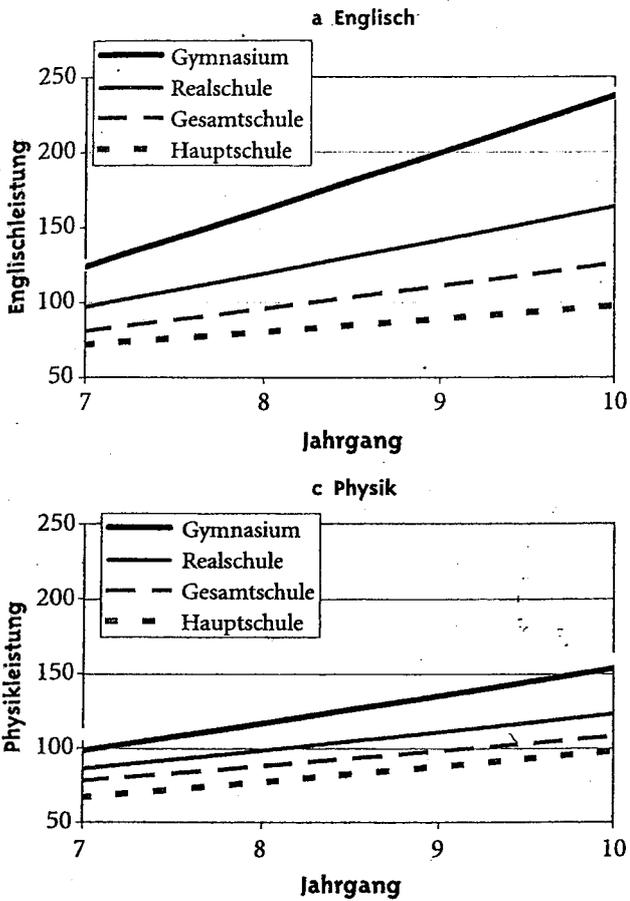


Abbildung 23.2. Entwicklung der Schulleistungen in drei Fächern von der 7. bis zur 10. Jahrgangsstufe nach Schulform

wurde ein sogenanntes Ankeritem-Design realisiert, bei dem wenigstens zu benachbarten Messzeitpunkten einige gemeinsame Items vorgegeben werden, um so mit Hilfe der entsprechenden Skalierungsverfahren einen einheitlichen Maßstab über die Zeit zu realisieren (zu den Details dieser Skalierungsprozeduren s. Köller, 1998). Die Metrik ist dabei so gewählt, dass zum ersten Erhebungszeitpunkt die Standardabweichung  $SD = 30$  in allen Fächern betrug. Erkennbar ist, dass die Leistungsentwicklungen in Englisch (als erster Fremdsprache) und Mathematik deutlich steiler verlaufen als in Physik, das weniger kontinuierlich und mit geringeren Stundenzahlen unterrichtet wird.

**Ergebnisse der BIJU-Studie.** Die Zuwächse für Gymnasiasten betragen z.B. in Englisch ca. 120 Punkte (4 Standardabweichungen), in Mathematik ca. 100 Punkte (3.3 Standardabweichungen), in Physik aber nur 50 Punkte (1.7 Standardabweichungen). Erschwerend zu den geringeren

Lerngelegenheiten in diesem Fach kann hier vermutet werden, dass die verschiedenen Stoffgebiete der Physik stark kompartimentalisiert sind, so dass sich in den niedrigen Lernraten über die Zeit auch eine fehlende Anknüpfbarkeit zwischen Vorwissen und neu hinzukommenden Inhalten widerspiegeln kann. Weiterhin ist in Abbildung 23.2 sichtbar, dass die „Leistungsschere“ zwischen den Schulformen vor allem zugunsten der Gymnasiasten über die Zeit zunehmend aufgeht. Inwieweit es sich bei diesem „Matthäus-Effekt“ („Wer hat, dem wird gegeben“), um systematische Effekte der Schulform handelt, wird unten ausgeführt.

## 2.2 Definition von schulischen Leistungsniveaus

Die in der Abbildung 23.2 gezeigten Entwicklungsverläufe von Schulleistungen belegen einen kontinuierlichen Anstieg über die Zeit, wobei

offenbleibt, welche konkreten Veränderungen in den Gedächtnisinhalten und kognitiven Operationen zu den quantitativen Sprüngen in den Leistungen führen. In diesem Abschnitt soll ein Vorgehen beschrieben werden, mit dem auf der Basis vorhandener Schülerleistungen in standardisierten Tests Rückschlüsse auf unterschiedliche Fähigkeitsniveaus möglich sind. Wiederum ist es die enge Anbindung von Schulleistungen an Curricula, die dazu führt, dass hier eine deutlich andere Forschungsstrategie gewählt wird, als es in entwicklungspsychologischen Arbeiten zur kognitiven Entwicklung üblich ist.

**Entwicklung als Reifungsprozess.** In den Bereichen der allgemeinen kognitiven Entwicklung, des Erwerbs von Problemlösekompetenzen, des Spracherwerbs oder auch der elementaren mathematischen Kompetenzen werden in entwicklungspsychologischen Theorien und Modellen die in unterschiedlichen Altersstufen erreichten Entwicklungsniveaus primär auf Reifungsprozesse zurückgeführt. Ein Beispiel für diese Form der Theorienbildung ist das Stufenmodell von Piaget zur kognitiven Entwicklung in der Kindheit und Jugend (vgl. den Beitrag von Montada in diesem Band), in dem für verschiedene Altersstufen unterschiedliche Fähigkeitsniveaus angenommen werden, deren Erreichung bzw. Nicht-Erreichung das Resultat von differenziellen Entwicklungsprozessen ist und sich mit Hilfe entsprechender Aufgaben diagnostizieren lässt. Im Falle von curricular verankerten Schulleistungen sind es die Lehrpläne (Curricula), die im Zusammenspiel mit der kognitiven Entwicklung definieren, in welchem Alter bzw. in welcher Jahrgangsstufe welche fachlichen Kompetenzen erreicht werden sollten. Für den Bereich der Mathematik sollte beispielsweise jeder Schüler bzw. jede Schülerin am Ende der Sekundarstufe I in der Lage sein, Aufgaben zur Prozentrechnung mit hinreichender Sicherheit zu lösen oder Volumina von Würfeln und Quadern berechnen können.

**Variabilität in der kognitiven Entwicklung.** Wie die Arbeiten zur kognitiven Entwicklung auf der Grundlage der Piagetschen Theorie gezeigt

haben (s. den Beitrag von Montada in diesem Band), findet man bei Personen des selben Alters eine erhebliche Varianz des kognitiven Entwicklungsniveaus. Keineswegs erreichen alle Jugendlichen das formal-operative Niveau. Das Gleiche gilt für die Schulleistungen; längst nicht alle Schülerinnen und Schüler erreichen in allen Fächern am Ende der Sekundarstufe I ein Leistungsniveau, wie es durch die Lehrpläne vorgegeben wird. Die Feststellung dieser Unterschiede in der individuellen Schulleistungsentwicklung erfolgt üblicherweise *post hoc* mit Hilfe der so genannten Skalenverankerungstechnik (*scale anchoring*; Beaton & Allen, 1992), auf die im Folgenden eingegangen werden soll.

**Skalenverankerung.** Neuere Schulleistungsuntersuchungen nutzen die psychometrischen Vorteile, die Tests bieten, die auf der Grundlage von *Item-Response-Modellen* (vgl. z.B. Hambleton & Swaminathan, 1989) konstruiert wurden. In deutschen Arbeiten wird üblicherweise anstelle von *Item-Response-Modellen* von der probabilistischen Testtheorie gesprochen (vgl. z.B. Rost, 1996). Solche Tests haben die Eigenschaft, dass Personenfähigkeiten und Itemschwierigkeiten auf derselben Dimension (dem selben Maßstab) abgetragen werden können. Hat also eine Person einen Fähigkeitswert von 500 – der Maßstab ist hier beliebig – und einige Items haben ebenfalls die Schwierigkeit 500, so heißt dies, dass die Person diese Aufgaben mit einer hinreichenden Sicherheit löst. (In den aktuellen Schulleistungsstudien wird unter hinreichender Sicherheit eine Lösungswahrscheinlichkeit von  $p = .65$  verstanden.) Aufgaben mit Schwierigkeitswerten deutlich unter 500 werden von dieser Person mit noch größerer Sicherheit gelöst, Aufgaben mit Schwierigkeitswerten deutlich über 500 werden mit großer Wahrscheinlichkeit von dieser Person nicht mehr gelöst. Anhand einer Analyse der Aufgabeninhalte mit einem Schwierigkeitswert von 500 kann dann festgestellt werden, welche kognitiven Operationen nötig sind, um die Items erfolgreich zu lösen. Man kommt so zu einer verhaltensnahen und inhaltsbezogenen Beschreibung der Ope-

rationen, die Personen mit einem Fähigkeitsparameter von 500 sicher beherrschen (zu Details des methodischen Vorgehens vgl. Klieme, 2000, und Klieme et al., 2000). Diese Art der Festlegung unterschiedlicher Leistungsniveaus soll genauer an den Aufgaben zur mathematischen Grundbildung illustriert werden, die im Rahmen der TIMS-Studie (Third International Mathematics and Science Study) eingesetzt wurden (vgl. Baumert et al., 2000a). In dieser Untersuchung wurden Schülerinnen und Schüler am Ende der Pflichtschulzeit, d.h. Gymnasiasten im letzten Jahr der gymnasialen Oberstufe und Auszubildende im letzten Jahr der beruflichen Erstausbildung, getestet. Alle Aufgaben thematisierten Stoffe aus den Lehrplänen der Sekundarstufe I und zeichneten sich zusätzlich dadurch aus, dass es um die Anwendung der mathematischen Kenntnisse auf alltagsnahe Probleme geht. Die Leistungen der Personen und Schwierigkeiten der Aufgaben wurden auf einem Maßstab mit einem Mittelwert von  $M=500$  und einer Standardabweichung von  $SD=100$  abgetragen. Die Abbildung 23.3 zeigt Aufgabenbeispiele mit ihren Schwierigkeiten, die im Sinne des „*scale anchorings*“ unterschiedliche Fähigkeitsniveaus definieren (s. Klieme et al., 2000, S. 122ff).

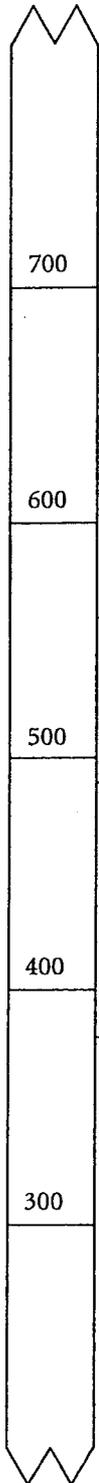
**Alltagsnahes Schlussfolgern.** Die erste Kompetenzstufe (Testwert von 400) umfasst Anforderungen, die sich im Wesentlichen auf das alltagsbezogene Schlussfolgern beschränken und keine expliziten mathematischen Operationen beinhalten. Dies gilt für die Aufgabe C6 in Abbildung 23.3 (S. 764), deren erfolgreiche Lösung lediglich eine einfache Überlegung erfordert: Je mehr Schritte jemand benötigt, um eine bestimmte Entfernung zu überwinden, desto kleiner ist seine Schrittlänge. Formal steht hinter dieser Aufgabe eine umgekehrte Proportionalität; es ist aber nicht nötig, diese für die Lösung zu berücksichtigen. Klieme et al. (2000) haben dieses Leistungsniveau als *alltagsnahes Schlussfolgern* bezeichnet, das bereits in der Grundschule erreicht sein sollte. Bemerkenswerterweise ergaben die Analysen in TIMSS, dass über vier Prozent der Gymnasias-

ten und über 20 Prozent der Auszubildenden für einen praktischen Beruf nicht über dieses Niveau hinaus kommen. Dies weist deutlich auf den konzeptuellen Unterschied zwischen Intelligenz und Schulleistungen hin. Junge Erwachsene mit Intelligenzwerten auf dem Niveau von Grundschulern würde man als schwer entwicklungsverzögert bzw. kognitiv retardiert bezeichnen. Im Hinblick auf Schulleistungen kann man solche Interpretationen nicht ohne weiteres vornehmen, die Personen können hinsichtlich ihrer allgemeinen kognitiven Entwicklung völlig unauffällig sein, es aber versäumt haben, die Lerngelegenheiten im Kontext Schule zu nutzen.

**Einfache mathematische Routinen.** Die *zweite Kompetenzstufe* in Abbildung 23.3 (Testwert von 500) umfasst die *sichere Anwendung von einfachen mathematischen Routinen*, wie sie in der Sekundarstufe I erworben werden sollten (Prozentrechnungen oder Flächenberechnungen). In diesem Sinne erfordert die Aufgabe D13 in Abbildung 23.3 (S. 764) die Beherrschung der Addition und Prozentrechnung. In der TIMSS-Oberstufenuntersuchung kamen rund 30 Prozent der Schülerinnen und Schüler in gymnasialen Oberstufen und ca. 65 Prozent der Personen in einer praktischen beruflichen Ausbildung nicht über dieses Niveau hinaus.

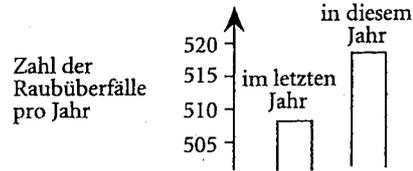
**Mathematisches Modellieren.** Die *dritte Kompetenzstufe* (Testwert von 600) erlaubt das *mathematische Modellieren auf einem einfachen Niveau*. Typische Aufgaben, die auf dieser Stufe gelöst werden können, erfordern die Verknüpfung unterschiedlicher Operationen und die Erschließung des mathematischen Ansatzes. In der Aufgabe B17 in Abbildung 23.3 (S. 764) müssen beispielsweise die Volumina beider Packungen berechnet werden und anschließend ist eine Verhältnisrechnung erforderlich. Kennzeichnend ist weiterhin, dass der mathematische Ansatz (hier: Volumenbestimmung und Verhältnisrechnung) nicht im Aufgabentext selbst nahe gelegt wird, sondern erschlossen werden muss. Auf der Basis der curricularen Vorgaben durch die Lehrpläne der Sekundarstufe I sollte erwartet werden, dass

Fähigkeit



**Aufgabe D17**

Ein Fernsehreporter zeigte dieses Diagramm und sagte:  
„In diesem Jahr hat die Zahl der Raubüberfälle stark zugenommen.“

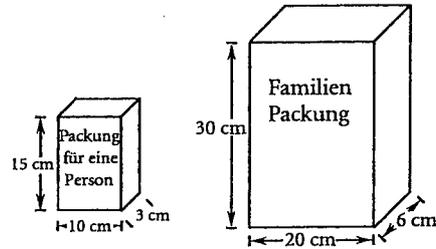


Halten Sie die Aussage des Reporters für eine angemessene Interpretation des Diagramms? Geben Sie eine kurze Erklärung!

681  
(0,1/0,19)

**Aufgabe B17**

Die beiden abgebildeten Müslipackungen haben die gleiche Form und sind ganz voll mit Müsli. Packung 1 enthält 80 Gramm Müsli.



Packung 1

Packung 2

Wieviel Gramm des gleichen Müsli enthält Packung 2?

- A. 160      B. 320      C. 480      D. 640

559  
(0,49/0,51)

488  
(0,64/0,64)

**Aufgabe D13**

Bei einer Wahl in einer Schule mit drei Kandidaten bekam Jan 120 Stimmen, Maria erhielt 50 Stimmen und Georg 30 Stimmen.

Welchen Prozentsatz der Gesamtstimmen bekam Jan?

- A. 60 %      B.  $66\frac{2}{3}$  %      C. 80 %      D. 120 %

380  
(0,83/0,86)

**Aufgabe C6**

Vier Kinder messen die Breite eines Zimmers. Sie zählen dabei, wie viele Schritte sie benötigen, um das Zimmer zu durchschreiten. Die Tabelle zeigt ihre Ergebnisse.

Name	Anzahl der Schritte
Stefan	10
Elke	8
Anna	9
Lars	7

Wer hat den längsten Schritt?

- A. Stefan      B. Elke      C. Anna      D. Lars

Die Werte an den Verbindungslinien zwischen den Beispielen und der Fähigkeitssäule geben das für eine 65-prozentige Lösungswahrscheinlichkeit erforderliche Fähigkeitsniveau und die Werte in Klammern die relativen internationalen und deutschen Lösungshäufigkeiten an. (IEA. Third International Mathematics and Science Study) © TIMSS/III-Germany

**Abbildung 23.3.** Testaufgaben aus dem Bereich mathematischer Grundbildung, die unterschiedliche Kompetenzniveaus repräsentieren

**Tabelle 23.1.** Beschreibung der unterschiedlichen Kompetenzstufen für drei Dimensionen des Leseverständnisses in der PISA-Studie (vgl. Baumert et al., 2001)

Dimension des Leseverständnisses			
	Informationen herausuchen	Textbezogenes Interpretieren	Reflektieren und Bewerten
<i>Aufgaben auf der jeweiligen Kompetenzstufe erfordern vom Leser ...</i>			
Stufe 5	... , verschiedene, tief eingebettete Informationen zu lokalisieren und geordnet wiederzugeben. Der Inhalt und die Form des Textes sind unbekannt; der Leser muss entnehmen, welche Information im Text für die Aufgabe relevant ist.	... ein vollständiges und detailliertes Verstehen eines Textes, dessen Format und Thema unbekannt sind.	... die kritische Bewertung oder das Bilden von Hypothesen, unter Zuhilfenahme von speziellem Wissen. Typischerweise verlangen Aufgaben dieses Niveaus vom Leser den Umgang mit Konzepten, die der Erwartung widersprechen.
Stufe 4	... , mehrere eingebettete Informationen zu lokalisieren. Üblicherweise sind der Inhalt und die Form des Textes unbekannt.	... z.B. das Auslegen der Bedeutung von Sprachnuancen in Teilen des Textes, die unter Berücksichtigung des Textes als Ganzes interpretiert werden müssen. Andere Aufgaben erfordern das Verstehen und Anwenden von Kategorien in einem unbekanntem Kontext.	... z.B. die kritische Bewertung eines Textes oder das Formulieren von Hypothesen über Information im Text, unter Zuhilfenahme von formalem oder allgemeinem Wissen. Leser müssen ein genaues Verstehen von langen und komplexen Texten unter Beweis stellen.
Stufe 3	... , Einzelinformationen herauszusuchen und dabei z.T. auch die Beziehungen dieser Einzelinformationen untereinander zu beachten. Die Auswahl wird durch auffallende und konkurrierende Informationen erschwert.	... , die in verschiedenen Teilen des Textes enthaltenen Aussagen zu berücksichtigen und zu integrieren, um eine Hauptidee zu erkennen, eine Beziehung zu verstehen oder die Bedeutung eines Wortes oder eines Satzes zu schlussfolgern. Beim Vergleichen, Kontrastieren oder Kategorisieren müssen viele Merkmale berücksichtigt werden.	... entweder Verbindungen, Vergleiche und Erklärungen oder sie erfordern vom Leser, bestimmte Merkmale des Textes zu bewerten. Genaues Verständnis des Textes im Verhältnis zu bekanntem Alltagswissen.
Stufe 2	... , eine oder mehrere Informationen zu lokalisieren, die beispielsweise aus dem Text geschlussfolgert werden müssen und mehrere Voraussetzungen erfüllen. Die Auswahl wird durch einige konkurrierende Informationen erschwert.	... z.B. das Erkennen des wenig auffallend formulierten Hauptgedankens eines Textes. Verstehen von Beziehungen oder Erfassen einer Bedeutung innerhalb eines Textteils auf der Basis von einfachen Schlussfolgerungen.	... z.B. einen Vergleich von mehreren Verbindungen zwischen dem Text und über den Text hinausgehendem Wissen. Bezugnahme auf persönliche Erfahrungen und Einstellungen um bestimmte Merkmale des Textes zu erklären.
Stufe 1	... , eine oder mehrere unabhängige, aber ausdrücklich angegebene Informationen zu lokalisieren. Üblicherweise gibt es eine einzige Voraussetzung, die von der betreffenden Information erfüllt sein muss, und es gibt, wenn überhaupt, nur wenig konkurrierende Informationen im Text.	... das Erkennen des Hauptgedankens des Textes oder der Intention des Autors bei Texten über bekannte Themen. Der Hauptgedanke ist dabei entweder durch Wiederholung oder durch früheres Erscheinen im Text auffallend formuliert.	... z.B., eine einfache Verbindung zwischen Information aus dem Text und weitverbreitetem Alltagswissen herzustellen. Der Leser wird ausdrücklich angewiesen, relevante Faktoren in der Aufgabe und im Text zu beachten.

große Teile der Schülerinnen und Schüler derartigen Aufgaben am Ende der Pflichtschulzeit lösen. Tatsächlich gelingt dies gut 70 Prozent der Gymnasiasten in der Oberstufe und rund 35 Prozent der Auszubildenden.

**Mathematisches Argumentieren.** Schließlich beinhaltet die *vierte und höchste Kompetenzstufe* (Testwerte von 700) das mathematische Argumentieren, insbesondere anhand von graphischen Darstellungen. In der Beispielaufgabe in Abbildung 23.3 (S. 764) muss der Bearbeiter erkennen, dass die dargestellte Zunahme der Zahl der Raubüberfälle mit zehn Fällen nur etwa 1/50 der Ausgangsbasis ausmacht und damit keinesfalls als starke Zunahme interpretiert werden darf. Dieses Fähigkeitsniveau erreichen 29 Prozent der Gymnasiasten und lediglich 3 Prozent der Auszubildenden. Zu bedenken ist hier noch einmal, dass es sich bei den untersuchten Personen um junge Erwachsene im Alter zwischen 18 und 21 Jahren handelt.

**Inhaltliche Interpretation von Entwicklungsverläufen.** Diese Technik des *scale anchorings* ist trotz des *post-hoc*-Vorgehens aus theoretischer Sicht von besonderer Bedeutung, da sie auch für die stärker inhaltliche Interpretation von Entwicklungsverläufen der Schulleistungen genutzt werden kann. Beispielsweise belegen die gerade vorggeführten Analysen der TIMSS-Daten, dass ein Leistungsunterschied von einer Standardabweichung (= 100 Punkte) in der Tat qualitativ unterschiedliche Kompetenzniveaus abbildet. Bedenkt man, dass die Gymnasiasten in der Abbildung 23.2b in Mathematik im Laufe von vier Schuljahren über drei Standardabweichungen gewinnen, so wird klar, dass es im Laufe dieser Zeit wiederholt zu deutlichen Veränderungen im Grundwissen und in der Beherrschung kognitiver Operationen gekommen sein muss. Im Rahmen der TIMS-Studie wurden vergleichbare Kompetenzstufen für die Naturwissenschaften in der Sekundarstufe I (vgl. Baumert, Lehmann et al., 1997) und der Sekundarstufe II (vgl. Baumert et al., 2000a, 2000b) definiert. Die PISA-Studie (*Programme for International Student Assessment*; vgl. Baumert, Artelt, Klieme & Stanat, 2001) liefert

darüber hinaus Kompetenzstufen für den Bereich des Textverständnisses. Hier wurden für die drei Subdimensionen *Informationen heraussuchen*, *textbezogenes Interpretieren* und *Reflektieren und Bewerten* die in Tabelle 23.1 (S. 765) aufgeführten Kompetenzstufen *post hoc* definiert.

! Insgesamt sollte dieser Abschnitt deutlich gemacht haben, dass erfolgreiche Schulleistungsverläufe in solchen Fächern zu erwarten sind, die kontinuierlich und kumulativ unterrichtet werden und dass *post-hoc*-Analysen der Schülerergebnisse in entsprechenden standardisierten Tests wichtige Aufschlüsse über unterschiedliche Kompetenz- bzw. Fähigkeitsniveaus liefern können.

## 3 Fähigkeitsgruppierungen und Schulleistungsentwicklung

### 3.1 Auswirkungen der Leistungsgruppierung

Aus einer erziehungswissenschaftlichen Sicht, aber auch aus einer entwicklungspsychologischen Perspektive, in der individuelle Veränderungen auf institutionelle Variablen zurückgeführt werden, ist es ohne Frage von Interesse, den Einfluss sich ändernder Schulumwelten auf die Entwicklung von Schulleistungen und anderen Persönlichkeitsmerkmalen zu untersuchen. Entsprechend dieser Sichtweise soll im folgenden Abschnitt auf die Effekte von Fähigkeitsgruppierungen im Bereich der Sekundarstufe eingegangen werden.

Charakteristisch für die Schulsysteme aller modernen Industrienationen ist die Trennung in einen Primar- bzw. Elementarbereich (die Grundschule) und einen Sekundarbereich. (Der tertiäre Bereich umfasst üblicherweise die universitäre und berufliche Ausbildung.) Beide Bereiche unterscheiden sich nicht nur in den Alters- bzw. Jahrgangsstufen ihrer Schülerinnen und Schüler, sie sind in der Regel auch geographisch getrennt.

Mit dem Übergang vom Primar- in den Sekundarbereich ist somit ein Ortswechsel und oftmals auch ein Wechsel der Lerngruppe verbunden.

### Leistungsgruppierung

Während der Grundschulbereich sich national und international dadurch auszeichnet, dass der Unterricht in sozial und leistungsmäßig unausgelesenen (heterogenen) Lerngruppen abläuft, findet in vielen Ländern nach dem Übergang in den Sekundarbereich früher oder später eine Leistungsgruppierung statt (vgl. hierzu Baumert et al., 2000a, S. 301ff). Schülerinnen und Schüler werden aufgrund ihrer Fachleistungen unterschiedlichen Schulformen zugewiesen, d.h. besonders leistungsstarke Kinder und Jugendliche besuchen dann einen gemeinsamen Schultypus, schwächere einen anderen. Der Grad der Differenzierung schwankt dabei national und international erheblich wie auch die Jahrgangsstufe, in der die Differenzierung beginnt. Während im deutschen Schulsystem die Zuweisung zu unterschiedlichen Schulformen zu Beginn der 5. Jahrgangsstufe erfolgt – Ausnahmen bilden die Bundesländer Berlin, Brandenburg und Bremen, wo der Sekundarbereich mit der 7. Jahrgangsstufe beginnt – führen die meisten anderen Nationen die Fähigkeitsgruppierungen später ein: In den Niederlanden beispielsweise nach der 6. Jahrgangsstufe, in Dänemark, Schweden oder Norwegen gar erst nach der 9. Jahrgangsstufe. In Australien und den USA wird oftmals völlig auf eine explizite Differenzierung in unterschiedliche Schulformen verzichtet. Den Differenzierungsmaßnahmen ist gemeinsam, dass die leistungsstarken Schülerinnen und Schüler auf einen Zweig wechseln, dessen erfolgreicher Abschluss sie zum Studium berechtigt, wohingegen die schwächeren Kinder und Jugendlichen dann Schulformen besuchen, die üblicherweise in eine berufliche Ausbildung münden.

**Differenzierung nach Fähigkeit.** Hinsichtlich der Fähigkeitsgruppierung werden im wesentlichen zwei Formen unterschieden (vgl. Arbeitsgruppe Bildungsbericht am MPI für Bildungsfor-

schung, 1994; Fend, 1982, 1991; Reuman, 1989), die externe Differenzierung (*between-classroom grouping*) und die Binnendifferenzierung (*within-classroom grouping*). Das gegliederte bundesdeutsche Schulsystem mit der Hauptschule, der Realschule und dem Gymnasium bietet ein Beispiel für die externe Differenzierung. Schülerinnen und Schüler dieser unterschiedlichen Schulformen besuchen getrennte Schulen, der Unterricht findet in leistungshomogenisierten Gruppen statt. Sonderformen der externen Differenzierung stellen die Kurssysteme der gymnasialen Oberstufe und der Integrierten Gesamtschulen dar, wie auch die unterschiedlichen Zweige der amerikanischen *high schools*. In den Gesamtschulen wird innerhalb einer Klasse in den Kernfächern (Deutsch, Mathematik, Fremdsprachen, Physik) ab einer bestimmten Jahrgangsstufe in zwei oder drei Kurse mit unterschiedlichen Leistungsniveaus differenziert, d.h. die Schüler einer Klasse werden in einigen Fächern gemeinsam, in anderen räumlich getrennt unterrichtet. Die Zuweisung zu den unterschiedlichen Kursniveaus basiert auf den Leistungen in diesen Fächern. In der gymnasialen Oberstufe, wie sie aktuell in allen Bundesländern mit Ausnahme Baden-Württembergs praktiziert wird, können die Jugendlichen zwei oder drei Leistungskurse wählen, die übrigen Fächer werden als Grundkurse unterrichtet. Abweichend von den üblichen Differenzierungen sind Leistungskurswahlen oftmals Neigungswahlen (vgl. Baumert & Köller, 2000), die der Schüler bzw. die Schülerin im Rahmen des Kursangebots der Schule autonom trifft.

**Das amerikanische Schulsystem.** In den USA werden in vielen Schulen unter einem gemeinsamen Dach drei unterschiedliche Zweige eingeteilt. Den *academic track* besuchen die leistungsstärksten Schülerinnen und Schüler, die später üblicherweise an eine Universität wechseln. Jugendliche des *vocational tracks* weisen schwächere Leistungen auf und wechseln in der Regel nach Abschluss der Schule in den Beruf. Schließlich besucht eine mittlere Gruppe den *general track*, ein Teil dieser Jugendlichen wechselt am Ende der

Schulzeit dann zum *College* oder einer Universität, der Rest geht in den Beruf über.

Binnendifferenzierung beschreibt im Gegensatz zur externen Differenzierung ein Vorgehen der Lehrkräfte innerhalb einer Schule bzw. Klasse, bei dem Schülerinnen und Schüler ohne geographische Trennung passend zu ihren unterschiedlichen Leistungs- bzw. Fähigkeitsniveaus Lerngruppen zugeordnet werden, in denen sie zum Teil spezifische Aufgaben und Bewertungen erhalten. Dieses stärker mikroadaptive Vorgehen innerhalb eines Klassenraumes verlangt von der Lehrkraft einen flexiblen Umgang mit der Leistungsheterogenität. Im deutschen Sekundarschulsystem findet man die Binnendifferenzierung sehr selten, da es hierzu der besonderen Zustimmung durch das entsprechende Kultusministerium bedarf, der Normalfall ist die externe Differenzierung.

#### Lernen leistungshomogene Gruppen besser?

Eine wichtige Ausgangsannahme aller externen Differenzierungsmaßnahmen ist, dass individuelle Lernerfolge in leistungshomogenen Gruppen höher sind als in leistungsheterogenen Gruppen (vgl. Arbeitsgruppe Bildungsbericht am Max-Planck-Institut für Bildungsforschung, 1994; Helmke & Weinert, 1997; Köller & Baumert, 2001). Autoren wie Carroll (1963) oder Bloom (1976) argumentieren in ihren Modellen schulischen Lernens, dass Lern- und Leistungsunterschiede zwischen Personen oftmals darauf beruhen, dass diese Personen unterschiedliche Zeiten benötigen, um den Unterrichtsstoff zu verstehen und abzuspeichern. Homogene Leistungsgruppen, so lässt sich aus diesen Modellen ableiten, benötigen ähnliche Lernzeiten, und gerade für leistungsstarke Gruppen impliziert dies, dass ein deutlich höheres Unterrichtstempo und kognitives Anspruchsniveau gewählt werden kann. Helmke und Weinert (1997, S. 145) konnten zeigen, dass die Vorkenntnisse der Schülerinnen und Schüler positive Effekte auf die Variablen „Klarheit des Unterrichts“ und „Zeitnutzung“ hatten, Lehrer also im Sinne makroadaptiven Handelns ihren Unterricht an das Leistungsniveau ihrer Klasse

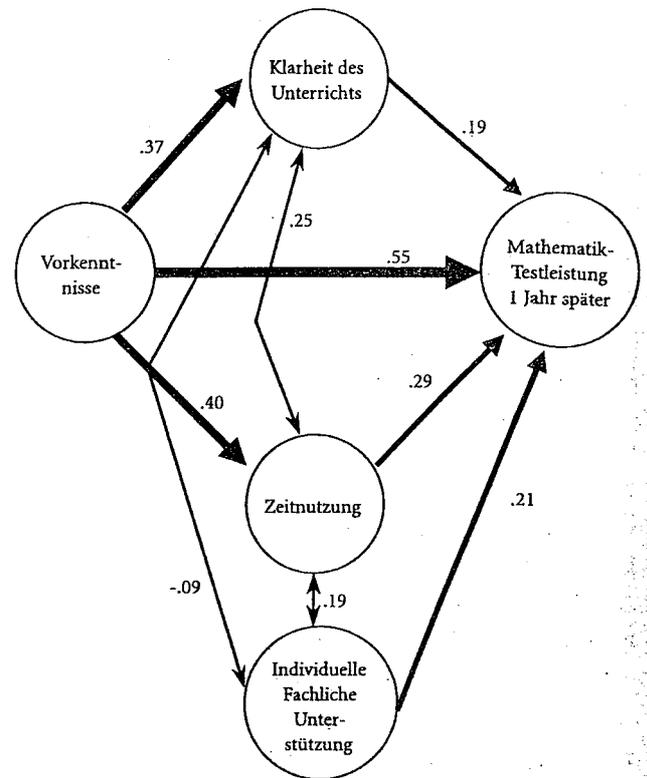


Abbildung 23.4. Leistungsstärke der Lerngruppe (Schulklasse), Unterrichtsmerkmale und Schulleistungsverlauf; standardisierte Pfadkoeffizienten aus Analysen auf Klassenebene (vgl. Helmke & Weinert, 1997, S. 145)

anpassen. In Folge dieses Anpassungsprozesses kam es zu positiven Effekten der Unterrichtsvariablen auf die Leistungsentwicklung der Klasse.

Das voll standardisierte Pfadmodell zeigt die Abbildung 23.4. Dort ist der Effekt des Vorkenntnisniveaus der Klasse auf die späteren Leistungen (.55) dokumentiert, den man etwas leichtfertig als Beleg für die Effizienz von Fähigkeitsgruppierungen interpretieren könnte, da Klassen mit höheren Vorkenntnissen später auch höhere Leistungen zeigen. In diesem Effekt sind aber leider die Einflüsse der individuellen und der auf Klassenebene gemittelten Vorkenntnisse konfundiert, so dass man den Koeffizienten von .55 nicht wirklich interpretieren kann.

Eine ähnlich elaborierte Studie wie die von Marsh et al. (s. „Unter der Lupe“, S. 769) haben Yang et al. (1999) publiziert. In die berichteten Analysen gingen rund 6400 Schülerinnen und

### Wie wirkt sich Leistungsgruppierung auf das Lernen aus?

Eine methodisch anspruchsvolle internationale Untersuchung, die sich mit den Effekten der Leistungsgruppierung auf die Entwicklung von Schulleistungen und selbstbezogenen Kognitionen bei Jugendlichen der Sekundarstufe I beschäftigt, wurde von Marsh et al. (2000) für das Schulsystem in Hongkong publiziert. Diese Studie überwindet die gerade skizzierten methodischen Probleme. Hongkong verfügt über ein besonders stark segregiertes Schulsystem, das heißt am Ende der 6. Jahrgangsstufe muss von allen Schülern ein landesweit verbindlicher Schulleistungstest bearbeitet werden, dessen Ergebnis als Kriterium dafür dient, welche *high school* den einzelnen Schüler anschließend aufnimmt. Besonders renommierte Schulen rekrutieren nur Schüler mit exzellenten Leistungen in den Tests, wohingegen prestigärere Schulen solche aufnehmen, die am Ende der Klassenstufe 6 schwächer abgeschnitten haben. Die unterschiedlichen *high schools* verfügen dementsprechend über sehr leistungshomogene Schülerschaften und die Zugehörigkeit zu den prestigeträchtigen Schulen gilt als Garant für eine erfolgreiche akademische bzw. berufliche Karriere. Durch diese Organisation des Bildungssystems lassen sich Effekte der Leistungsgruppierung optimal untersuchen.

In der konkreten Untersuchung von Marsh et al. (2000) wurden längsschnittliche Datensätze von rund 8000 Schülern aus 44 Schulen berücksichtigt. Als Prädiktoren für die Vorhersage der Schulleistungen in den Fächern Chine-

sisch (Muttersprache), Mathematik und Englisch, die in der 7., 8. und 9. Jahrgangsstufe erhoben wurden, dienten die individuelle Ausgangsfähigkeit zu Beginn der 7. Jahrgangsstufe und die auf Schulebene aggregierten Schulleistungen. Dieses aggregierte Maß bildet unmittelbar die Fähigkeitsgruppierung ab: je höher der aggregierte Wert, desto stärker das Leistungsniveau der Schule. Die von Marsh et al. berichteten Ergebnisse weisen auf eine unbedeutende Rolle der Fähigkeitsgruppierung hin. Beispielsweise ergab sich bei Kontrolle der individuellen Ausgangsleistung ein nicht signifikantes Regressionsgewicht ( $\beta = .10$ ) von der auf Schulebene aggregierten Ausgangsleistung auf die individuelle Leistung in der 9. Jahrgangsstufe. Ein deutlicher Effekt ( $\beta = .53$ ,  $p < .01$ ), ergab sich dagegen von der individuellen Ausgangsleistung auf die Leistung in der 9. Jahrgangsstufe. Die Ergebnisse lassen sich dahingehend interpretieren, dass sich die Schulleistungen guter Schüler in Hongkong unabhängig von der Leistungsstärke (der Fähigkeitsgruppierung) ihrer Schule positiver entwickeln als die Leistungen schlechter Schüler, d.h. hier zeigte sich der bereits erwähnte *Matthäus-Effekt* auf Individualebene, nicht aber auf Schulebene. Bemerkenswert an dieser Untersuchung ist ohne Frage, dass man bei simultaner Modellierung der Ausgangsleistung auf Individual- und Schulebene zu gänzlich anderen Befunden kommen kann als es der Fall ist, wenn die Analysen nur auf Schul- oder Individualebene durchgeführt werden.

Schüler aus 161 Schulen ein, deren Leistungen in Mathematik, Lesen, Schreiben und Naturwissenschaft längsschnittlich untersucht wurden. Für die auf Schulebene aggregierten Ausgangsleistungen in Mathematik und Naturwissenschaft ergaben

sich bei Kontrolle der individuellen Ausgangsleistungen sogar durchgängig negative Effekte auf die späteren Leistungen in allen vier untersuchten Domänen, d.h. je höher die Leistungen in beiden Fächern in einer Schule lagen, desto ungünstiger

war bei Kontrolle der individuellen Ausgangsleistungen die Wissensentwicklung.

**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 Deutschland 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. Jahrgangsstufe in Deutsch, Mathematik und Englisch aus einem Bundesland mit sechsjähriger Grundschule mit denen der übrigen Bundesländer mit vierjähriger Grundschule verglichen. Die Tabelle 23.2 zeigt die Resultate der Studie.

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

	Englisch	Mathematik	Deutsch	N
sechsjährige	$\bar{x}$ 38.8	38.0	49.5	1049
Grundschule	<i>s</i> 9.4	11.9	11.6	
vierjährige	$\bar{x}$ 52.6	47.1	54.8	11.203
Grundschule	<i>s</i> 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 frühere 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 beschriebenen Arbeiten von Marsh et al. (2000) und Yang et al. (1999) einen mehrerebenenanalytischen Auswertungsansatz. Basierend auf einer Stichprobe von  $N = 2730$  Schülerinnen und Schülern aus 107 Schulen 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. Die Analysen zeigten das folgende Ergebnismuster. Bei Kontrolle der übrigen Prädiktoren hatte die individuelle Ausgangsfähigkeit in Klasse 7 einen deutlich positiven Effekt auf die Leistung in Klasse 10 ( $\beta = .50$ ). 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 ein nur schwacher positiver Effekt ( $\beta = .10$ , *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 makro-adaptiven Verhaltens der Lehrkräfte auf die vorgefundene Leistungsstärke in den jeweiligen Gymnasialklassen, 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.

### 3.2 Kosten der Leistungsdifferenzierung im Sekundarbereich

Verschiedene Autoren haben die besonderen psychologischen Kosten der frühen Leistungsgruppierungen im Sekundarbereich hervorgehoben, wobei die Kritik oftmals stärker ideologisch und weniger theoretisch abgeleitet oder empirisch fundiert war. Zu solchen Kritikpunkten zählen (vgl. Arbeitsgruppe Bildungsbericht, 1994):

- ▶ Die unterschiedlichen Schulformen sind zu scharf gegeneinander abgegrenzt, so dass mit dem Übertritt von der Grundschule zur Sekundarschule oftmals die Weichen für das Leben des Einzelnen zu starr gestellt werden.
- ▶ Die Durchlässigkeit des Systems ist zu gering; nur wenige Schüler wechseln nach dem Übergang mit zehn oder elf Jahren noch einmal die Schulform. Wenn überhaupt, so geschieht dies quasi nur in Form von Abstufungen auf ein niedrigeres Niveau.
- ▶ Da viele Berufslaufbahnen in der Bundesrepublik Deutschland an formale Schul- bzw. Hochschulabschlüsse gekoppelt sind, bedeutet die Übergangsentscheidung auf eine der drei

Schulformen, dass bereits im Alter von zehn oder elf Jahren eine folgenschwere Entscheidung für das weitere Berufsleben der Kinder gefällt wird.

- ▶ Das differenzierte Schulsystem benachteiligt Kinder und Jugendliche aus sozial schwächeren Schichten, die überproportional häufig die Hauptschule besuchen und damit geringere Karrierechancen haben.

Jenseits dieser sozialkritischen Beleuchtung des differenzierten Schulsystems hat sich auch die psychologische Forschung mit den Kosten der Fähigkeitsgruppierungen beschäftigt. Aus entwicklungspsychologischer Sicht ist diese Thematik von besonderem Interesse, da die Fähigkeitsgruppierung in der Regel mit dem Aufkommen der Pubertät zusammenfällt, in der weitreichende biologische, kognitive und psychosoziale Veränderungen stattfinden. Zwei psychologische Ansätze, die sich mit den individuellen Schwierigkeiten beim Übergang von der unselegierten Grundschule in den segregierten Sekundarbereich beschäftigen, sollen im Folgenden geschildert werden.

#### Das Developmental/Stage Environment Fit Model

Der Übergang, vom Grundschul- in den Sekundarbereich findet in der Regel im Alter von 11 bis 13 Jahren statt, jenem Lebensabschnitt also, in dem sich in Folge der einsetzenden Pubertät massive Veränderungen auf Seiten der Jugendlichen einstellen. In einer Vielzahl von Studien (s. im Überblick Wigfield et al., 1996) wurden die Entwicklungsverläufe von Jugendlichen in dieser Zeit des Schulwechsels untersucht. Dabei zeigten sich die folgenden Befunde:

- ▶ Mit dem Übergang von der Grundschule in den Sekundarbereich verschlechtern sich oftmals die Schulleistungen (Simmons & Blyth, 1987).
- ▶ Gleichzeitig kommt es zu einem Absinken der intrinsischen Lernmotivation (Gottfried et al., 2001; Harter, 1981).
- ▶ Die Leistungsangst steigt dagegen an (Hill, 1980).

- ▶ Die *drop-out-Raten*, d.h. die Zahl der Schüler, die aus dem Schulsystem aussteigen, nehmen deutlich zu (Vallerand et al., 1997).

Eccles und Mitarbeiterinnen (z. B. Eccles & Midgley, 1989; Eccles et al., 1993) haben auf der Grundlage dieser Befunde das so genannte *Developmental/Stage Environment Fit Model* vorgestellt, das versucht, diese unerwünschten Entwicklungsläufe als Folge einer fehlenden Passung zwischen individuellen Bedürfnissen in der Pubertät und den schulischen Rahmenbedingungen zu konzeptualisieren. Die zentralen Modellannahmen sind dabei:

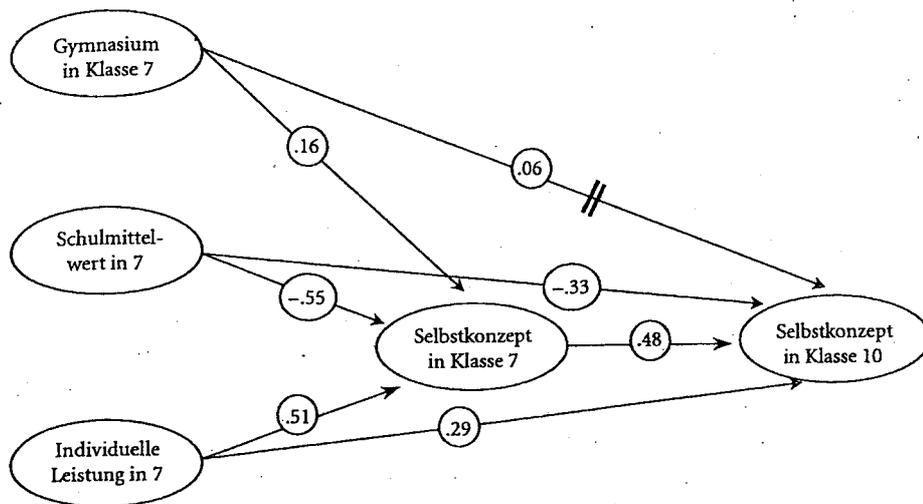
- ▶ Umweltbedingungen, die inkompatibel zu individuellen Bedürfnissen sind, haben negative Effekte auf subjektive Wohlbefindensmaße, motivationale Variablen und leistungsthematisches Verhalten.
- ▶ Mit der einsetzenden Pubertät steigt das individuelle Bedürfnis nach Autonomie und Identitätsbildung.
- ▶ Schulen bzw. Lehrkräften gelingt es in dieser Phase nicht, die Balance zwischen den Bedürfnissen der Schüler nach Autonomie und den institutionellen, einschränkenden Vorgaben herzustellen (vgl. z.B. Hunt, 1975).
- ▶ Zudem ist der deutliche Anstieg des kompetitiven Klassenklimas in Sekundarschulen oftmals inkompatibel zum individuellen Bedürfnis nach Kompetenzerleben (Ames, 1992; Covington, 1992).
- ▶ Die Folge sind negative Effekte auf Schulleistungen und die intrinsische Lernmotivation (im Überblick Wigfield et al., 1996).

Empirische Befunde zur Stützung des Modells findet man bei Eccles und Midgley (1989) oder auch in der Überblicksarbeit von Wigfield et al. (1996). Danach besteht kein Zweifel daran, dass eine Lernumwelt, die stärker an den Bedürfnissen der Jugendlichen ausgerichtet ist, deutlich positive Effekte auf die Entwicklung leistungsbezogener Variablen wie Lernmotivation, Fähigkeitsselbstbilder etc. hat.

### Der Fischteich-Effekt

Insbesondere die Arbeiten von Schwarzer und Mitarbeitern (Schwarzer, 1979; Schwarzer & Jerusalem, 1982; Schwarzer et al., 1982) für das deutsche Schulsystem und von Marsh (z.B. Marsh, 1987, 1990, Marsh et al., 2000; Marsh & Parker, 1984) für das australische und amerikanische Schulsystem haben deutlich gemacht, dass die Fähigkeitsgruppierungen im Sekundarschulbereich deutliche Effekte auf die Entwicklung selbstbezogener Fähigkeitskognitionen haben. Für leistungsstarke Schüler hat der Übergang auf das Gymnasium für die Entwicklung ihrer selbst wahrgenommenen Fähigkeiten erhebliche negative Einflüsse. Gehörten diese Schüler in der Grundschule noch zu den Besten, so erleben sie auf dem Gymnasium, dass viele andere in der Leistung ebenbürtig oder besser sind. Soziale Vergleiche (Festinger, 1954) führen hier eher zu einem Absinken fähigkeitsbezogener Selbstkonzepte und des Selbstwertgefühls. Schwache Schülerinnen und Schüler erleben dagegen auf Hauptschulen, dass sie plötzlich nicht mehr die Leistungsschwächsten in der Klasse sind, was zu einem Anstieg ihrer Fähigkeitsselfkonzepte führt. Dieser Prozess mündet oft darin, dass das mittlere Fähigkeitsselfkonzept auf den verschiedenen Schulformen im Laufe der Sekundarstufe I konvergiert (z.B. Schwarzer & Jerusalem, 1982).

**Schwache Mitschüler stärken das Selbstwertgefühl.** Das Phänomen, wonach Leistungsgruppierungen Effekte auf selbstbezogene Fähigkeitskognitionen haben, hat Marsh (1987, 1990) in Anlehnung an Davis (1966) als *big-fish-little-pond-effect* beschrieben. Zwei Schüler oder Schülerinnen (*fishes*) mit gleicher individueller Leistungsfähigkeit, die aber Klassen bzw. Schulen mit unterschiedlichen Leistungsniveaus besuchen, weisen unterschiedliche Selbstwahrnehmungen eigener Fähigkeiten auf; das heißt, der Schüler (*big fish*) in der schwächeren Schule (*little pond*) hat eine höhere Wahrnehmung eigener Fähigkeiten als der entsprechende Schüler (*little fish*) in der leistungsstärkeren Schule (*big pond*). In empirischen Studien spiegelt sich dieser *Fischteich-*



**Abbildung 23.5.** Effekte der Leistungsgruppierung in der Sekundarstufe I auf die Entwicklung des Selbstkonzepts eigener Begabung im Fach Mathematik (Befunde aus Köller & Baumert, 2001)

Effekt in negativen Regressions- oder Pfadkoeffizienten der über die Schüler gemittelten Klassen- oder Schulleistung auf die individuelle Selbstwahrnehmung der Begabung wider, wenn der Einfluss der individuellen Leistung konstant gehalten wird. Ein entsprechendes Pfaddiagramm für das Selbstkonzept eigener Begabung in Mathematik zeigt die Abbildung 23.5. Die dargestellten Ergebnisse stammen aus der oben bereits beschriebenen Studie von Köller und Baumert (2001).

Der interessanteste Effekt in Abbildung 23.5 ist der Pfad von der auf Schulebene gemittelten Mathematikleistung in Klasse 7 auf das Selbstkonzept der Begabung in der 10. Jahrgangsstufe bei Kontrolle der Selbstkonzeptvariable in Klasse 7. Der Koeffizient von  $\beta = -.33$  kann direkt als Effekt der Leistungsgruppierung auf die Entwicklung des Selbstkonzepts der Begabung interpretiert werden. Je leistungsstärker die Klassen bzw. die Schulen sind, desto ungünstiger ist der Entwicklungsverlauf ihrer selbstbezogenen Fähigkeitskognitionen. Dieser Effekt zeigt sich im übrigen auch für schulische Interessen (vgl. Köller, Schnabel & Baumert, 2000).

**FAZIT**

Zusammenfassend belegt dieser Abschnitt, dass Änderungen der institutionellen Lernumwelten erhebliche Effekte auf die Leistungsentwicklung und Veränderung motivationaler und psychosozialer Merkmale haben können.

**4 Schulleistungen – Leistungen der Schule oder des Schülers?<sup>6</sup>**

Bis hier sind wir von der Voraussetzung ausgegangen, dass Schulen die entscheidende Instanz bei der Vermittlung zentraler Wissensinhalte sind. Schulleistungen und ihre Entwicklung, so die Annahme, sind unmittelbar an die Lerngelegenheiten im Unterricht geknüpft (vgl. Baumert & Köller, 1998). Rutter et al. (1979) schätzen, dass Schülerinnen und Schüler ca. 15.000 Stunden

<sup>6</sup> Diesen Titel hat der unvergessene Franz E. Weinert 2001 für einen seiner letzten Aufsätze gewählt. Wir möchten mit der Übernahme dieses Titels unsere Achtung und Anerkennung gegenüber dem Lebenswerk dieses großartigen Wissenschaftlers ausdrücken.

ihres jungen Lebens in Schulen verbringen, die Schule somit neben der Familie die wichtigste Entwicklungsumwelt in der Kindheit und Jugend ist. Entsprechend dieser besonderen Bedeutung der Schule als Entwicklungsmilieu investieren moderne Industrienationen einen erheblichen Anteil ihrer ökonomischen Ressourcen in die schulische Bildung, immer mit der impliziten oder expliziten Forderung, dass die Schule Wissenserwerbsprozesse auf Seiten ihrer Schülerinnen und Schüler initiiert, aufrecht erhält und somit erfolgreiche Bildungs- und Lebenskarrieren ermöglicht. Vor diesem Hintergrund ist es bemerkenswert, wie groß die Zweifel sind, die aus wissenschaftlicher Perspektive gegen die pädagogische Effizienz von Schulen hervorgebracht werden. In den 60er Jahren des letzten Jahrhunderts machte der so genannte *Coleman-Report* (Coleman et al., 1966) in den USA Furore. Basierend auf einer varianzanalytischen Auswertungsstrategie kamen die Autoren zu dem deprimierenden Ergebnis, dass nach Kontrolle von Schülermerkmalen lediglich fünf (bei weißen Schülern) bzw. neun (bei farbigen Schülern) Prozent der Leistungsunterschiede auf Schuleffekte zurückzuführen waren, 95 bzw. 91 Prozent der Unterschiede also durch andere Faktoren erklärt werden konnten. In den Sekundäranalysen von Jencks et al. (1973) „verdüsterte“ sich das Bild weiter, die Autoren schätzten jetzt, dass lediglich ein Prozent der Leistungsvarianz auf Schuleffekte in der Sekundarstufe zurückzuführen war. Jencks et al. kamen auf Grundlage ihrer Befunde zu folgenden Schlussfolgerungen (vgl. Weinert, 2001, S. 74):

- ▶ Genetische Unterschiede, die sich beispielsweise in der unterschiedlichen Intelligenzhöhe von Schülerinnen und Schülern manifestieren, erklären 33 bis 50 Prozent der Schulleistungsunterschiede.
- ▶ Unterschiedliche außerschulische Lernumwelten erklären 25 bis 40 Prozent der Leistungsunterschiede.
- ▶ Die unterschiedliche soziale Herkunft der Schülerinnen und Schüler erklärt ungefähr 6 Prozent der Leistungsvarianz.

- ▶ Unterschiede im zeitlichen Umfang der Schulbildung erklären 5 bis 15 Prozent der Leistungsvarianz.
- ▶ Differenzen in der Qualität der Grundschulen erklären 3 Prozent der Varianz und
- ▶ Qualitätsdifferenzen zwischen den Sekundarschulen erklären die oben bereits erwähnten ein Prozent der Leistungsunterschiede zwischen Personen.

**Einfluss der Schule auf Schülerleistungen.** Für Deutschland haben Ditton und Kreckler (1995) geschätzt, dass die Unterschiede zwischen Schulen innerhalb einer Schulart in einem Bundesland etwa 2,5 Prozent der Schülerdifferenzen erklären. Auch wenn neuere, methodisch elaboriertere Untersuchungen ein günstigeres Bild zeichnen – Scheerens und Bosker (1997) kommen beispielsweise aufgrund einer Meta-Analyse zu der Schätzung, dass es ca. 12 Prozent der Leistungsunterschiede zwischen Schülern sind, die auf Schuleffekte zurückführbar sind, Opdenakker und Van Damme (2000) wie auch Baumert und Köller (1998) präsentieren noch deutlich günstigere Zahlen – bleibt der Eindruck der begrenzten Effizienz der Bildungsinstitution Schule bestehen. Aus einer stärker lern-, gedächtnis- und entwicklungspsychologischen Perspektive haben die meisten dieser Studien aber die methodische Schwäche, dass es allein um die Erklärung interindividueller Differenzen geht, intraindividuelle Leistungsverläufe oder interindividuelle Unterschiede in der intraindividuellen Leistungsentwicklung jedoch unberücksichtigt bleiben. Interessante Fragen wie z.B.: „Wie würde die Leistungsentwicklung in den Fremdsprachen oder Naturwissenschaften in der Kindheit und Jugend verlaufen, wenn es keine institutionalisierten Bildungssysteme mit systematischer Instruktion gäbe?“ werden nicht gestellt, da sie wenigstens in modernen Industrienationen aufgrund der Schulpflicht auch kaum zu beantworten sind.

„In den hochindustrialisierten Ländern ist die entwicklungspsychologische und gesellschaftspolitische Bedeutung der Schule nicht so leicht

erkennbar, weil fast alle Kinder mehr als neun Jahre ihrer Schulpflicht genügen, weil die nationalen Schulsysteme durch die Ausbildung und Professionalisierung der Lehrer, durch eine staatlich geregelte Schulorganisation, durch materielle Mindestausstattung mit Lehr- und Lernmitteln sowie durch mehr oder minder verbindliche Lehrpläne und Prüfungsordnungen dafür gesorgt wird, dass im Prinzip mehr formale Ähnlichkeiten als spezifische Unterschiede zwischen den einzelnen Schulen und Schulklassen bestehen. Auf diese Weise verstärkt sich der Eindruck, dass es Unterschiede der Schüler und nicht Qualitätsdifferenzen der Schulen sind, welche die Stärken und Schwächen in den Schulleistungen bewirken.“ (Weinert, 2001, S. 79)

Die Schuleffizienzforschung hat somit das Problem, dass es bei der Untersuchung des Einflusses des *treatments* Schule bzw. Unterricht schlecht möglich ist, ein Kontrollgruppendesign zu wählen. Hätte man ein solches Design, würde man mit Sicherheit zu deutlich günstigeren Schätzungen für die Effektivität von Schulen kommen. Kaum jemand, der keine Schule besucht hätte, würde jemals tiefere Einblicke beispielsweise in die Quantenmechanik oder in die lateinische Sprache gewinnen, egal wie intelligent er ist und wie sein sozialer Hintergrund ist. Um eine Analogie aufzunehmen: Ein großer, nicht zu schwerer, muskulöser Mensch würde, wenn er einmal das Schwimmen gelernt hätte, immer schneller schwimmen als ein kleiner dicker, unmuskulöser Mensch. Für diese Leistungsunterschiede würden in hohem Maße die genetisch bedingten Differenzen in der Konstitution verantwortlich sein und nicht der Umstand, dass beide Personen vor langer Zeit einmal an einem Schwimmkurs teilgenommen hätten. Auf der anderen Seite würden beide ertrinken, d.h. Leistungsunterschiede wären nicht beobachtbar, wenn sie nicht zuvor diesen Schwimmkurs besucht hätten, in dem sie systematisch in die Schwimmtechnik eingewiesen worden wären. Übersehen wird also oftmals, dass

institutionalisierte Bildung im wesentlichen durch die Herstellung systematischer Lernsituationen charakterisiert werden kann, die notwendige Voraussetzungen für eine Veränderung der kognitiven Strukturen (nichts anderes ist Lernen, s.o.) darstellen, wiewohl interindividuelle Unterschiede im Ausmaß der Veränderung auftreten, die oftmals in der Tat auch Folge unterschiedlicher genetischer Ausstattung sein können.

**Strukturierter Unterricht kompensiert fehlende Gedächtnisstrategien.** Die Bedeutung der Schule für den Wissenserwerbsprozess kann sehr gut anhand konkreter Gedächtnisstrategien erläutert werden, die dazu dienen, neues Wissen abzuspeichern (vgl. hierzu den Beitrag von Schneider & Büttner in diesem Band). Der aktive selbstregulierte Einsatz solcher Wiederholungs-, Organisations-, Elaborations- und Abrufstrategien gelingt jüngeren Kindern kaum, erst im Laufe der Entwicklung steigt die Verfügbarkeit solcher Strategien. Die im schulischen Fachunterricht hergestellten Lernsituationen übernehmen quasi kompensatorisch die Rolle der bei den Schülerinnen und Schülern defizitär ausgeprägten Strategien zur Speicherung und zum Abrufen von Gedächtnisinformationen. In neueren Arbeiten (vgl. Köller et al., im Druck) wird dementsprechend auch Schulen die Bedeutung von *executive functionaires* zugesprochen, die dem kognitiven System der Kinder und Jugendlichen Teile der Organisation des Informationsverarbeitungsprozesses in Lernsituationen abnehmen. Folgt man diesem Ansatz, so lässt sich schlussfolgern, dass mit zunehmendem Alter und der fortschreitenden Ausbildung von individuellen Gedächtnisstrategien die Schule, besser: hoch strukturierte Unterrichtssituationen, für gelingende Wissenserwerbsprozesse unwichtiger werden. Dies ist auch in der Tat der Fall. Beispielsweise ist der Unterricht an Universitäten deutlich unstrukturierter und ein erfolgreiches Studium ist sehr viel stärker an selbstreguliertes Lernen gebunden als die Bewältigung der Anforderung im Grundschul- und Sekundar-schulbereich.

**Primäre und sekundäre kognitive Fähigkeiten.**

Die besondere Rolle der Schule für die Leistungsentwicklung in der Kindheit und Jugend betont auch der Evolutions- und Entwicklungspsychologe Geary (1995, 1996), der zwischen *primären* und *sekundären kognitiven Fähigkeiten* unterscheidet. Zu den primären Fähigkeiten zählt er die angeborenen Prädispositionen zum Erwerb der Muttersprache, des elementaren numerischen Verständnisses, der konkreten intellektuellen Operationen und des ökologisch bedeutsamen Weltwissens. „Diese Fähigkeiten sind offenbar genetisch und hirnganisch bei allen gesunden Menschen so prädisponiert, dass sie auch unter minimalen, oft sehr ungünstigen Umwelt-, Entwicklungs- und Bildungsbedingungen erworben werden“ (Weinert, 2001, S. 82). Für die Mathematik nennt Geary (1996) vier grundlegende angeborene Fähigkeiten:

- (1) Die Fähigkeit zur akkuraten Abschätzung kleiner Mengen, ohne die Elemente zählen zu müssen (*Numerosity*).
- (2) Ein grundsätzliches Verständnis von größer-kleiner-Relationen, sofern die Zahl der Elemente nicht 5 übersteigt (*Ordinality*).
- (3) Die Fähigkeit, ohne vorherigen Spracherwerb zu zählen (*Counting*). Geary (1996, S. 232) spricht hierbei vom „*preverbal counting system*“.
- (4) Das Addieren und Subtrahieren von kleinen Mengen (*Simple Arithmetic*).

Komplexere mathematische Kompetenzen, beispielsweise in der Geometrie oder Algebra, zählt Geary zu den sekundären Fähigkeiten, die kulturell vermittelt werden: „In addition, it is likely that most features of complex mathematical domains, such as algebra, geometry, and calculus are biologically secondary, given that the associated abilities only emerge with formal education“ (Geary, 1996, S. 233). *Formal education* meint hier nichts anderes als schulische Lehr-Lernprozesse.

**Schulleistungsuntersuchungen.** Relativ gute Hinweise für die hohe Bedeutung von Bildungssystemen für die individuelle Schulleistungsentwicklung geben die in letzter Zeit publizierten

internationalen Schulleistungsuntersuchungen (Baumert et al., 2000a, 2000b; Baumert, Lehmann et al., 1997). Der dort durchgeführte Vergleich zwischen Entwicklungsländern (z.B. Südafrika) und Industrienationen (z.B. Deutschland) offenbart für die mathematisch-naturwissenschaftlichen Leistungen Unterschiede in der Größenordnung von zwei und mehr Standardabweichungen. Gegen die Interpretierbarkeit solcher Differenzen mag man noch einwenden, dass Unterschiede in den Bildungssystemen mit sozialen und kulturellen Unterschieden zu stark konfundiert seien. Stringenter sind solche Gegenüberstellungen, in denen man die Leistungen der Schülerinnen und Schüler aus verschiedenen modernen Industrienationen vergleichen kann, und zwar unter der realistischen Annahme, dass die Verteilungen intellektueller und sozialer Merkmale zwischen diesen Ländern vergleichbar sind, Leistungsunterschiede also am ehesten durch unterschiedlich effiziente Schulsysteme zustande kommen dürften. Hier zeigen sich dann in der Tat Effektstärken, die selbst für experimentelle Psychologen beeindruckend sein dürften. So berichten Baumert, Bos und Watermann (2000, S. 141) die folgenden Befunde zu den Mathematikleistungen in gymnasialen Oberstufen:

„Vergleicht man in einem zweiten Schritt die jeweils teststärksten 10 Prozent . . . , so liegen die Leistungen deutscher und österreichischer Schüler im unteren Mittelbereich . . . Nur die Schüler aus Italien, Griechenland, Tschechien und den USA schneiden schlechter ab . . . Die testleistungsstärksten Schüler aus allen anderen Ländern erreichen signifikant bessere Testergebnisse. Die Differenzen zu den deutschen Ergebnissen liegen zwischen 25 Punkten (Schweden) und 90 Punkten (Slowenien) ( $d = .55$  bzw.  $d = 1.97$ ).“

Die berichtete Effektstärke von  $d = 1.97$  für den Leistungsvorsprung slowenischer gegenüber deutschen Schülern wird kaum jemand ernsthaft auf Unterschiede in den kognitiven Grundfähigkeiten oder motivationalen Variablen zurückführen wol-

len, hier schlagen offensichtlich die Unterschiede in nationalen Bildungssystemen auf die Individualebene durch.

#### FAZIT

Unterrichtsprozesse in Schulen sind eine notwendige Voraussetzung für den Erwerb und Ausbau komplexerer fachgebundener Kompetenzen.

## 5 Das Zusammenspiel der Entwicklungsverläufe von Schulleistungen, Interessen, Fähigkeitsselbstkonzepten und der Intelligenz

Die stärker pädagogisch-psychologisch orientierten Arbeiten zu Schulleistungen (vgl. Helmke & Weinert, 1997; Walberg, 1981) haben die Identifikation und Analyse von Schulleistungsdeterminanten in den Vordergrund ihrer Forschungsbemühungen gestellt. In Ergänzung zu dieser Perspektive wollen wir in diesem Kapitel für ausgewählte Variablen beschreiben, wie deren Entwicklung durch die Schulleistungsentwicklung beeinflusst wird und *vice versa*. Die berücksichtigten Variablen sind die Intelligenz, die oben bereits aufgegriffenen Fähigkeitsselbstkonzepte und die schulischen Interessen. Alle drei Variablen gelten als wichtige individuelle Prädiktoren leistungsthematischen Wahlverhaltens (z.B. Köller et al., 2000) und schulischen Lernens (Byrne, 1996; Helmke & Weinert, 1997; Köller, Schnabel & Baumert, 2001; Marsh & Yeung, 1997; Möller & Köller, 2001).

Neuere nationale Längsschnittstudien wie die bereits oben erwähnte SCHOLASTIK-Studie oder auch das BIJU-Projekt (s.o.) erlauben Analysen, inwieweit es zu einer wechselseitigen Beeinflussung von Schulleistungen und den drei oben genannten Merkmalen kommt.

### 5.1 Schulleistungsentwicklung und Intelligenzentwicklung

Die psychometrische Intelligenz gilt als einer der zentralen Prädiktoren für die Schulleistungsentwicklung in den Kernfächern (Sprachen, Mathematik, Naturwissenschaften und Sozialwissenschaften). Einfache Produkt-Moment-Korrelationen zwischen Maßen der allgemeinen Intelligenz und Schulleistungsindikatoren liegen im Mittel bei ungefähr  $r = .50$ , allerdings mit erheblichen Schwankungen (zwischen  $.10$  und  $.90$ ; vgl. Helmke, 1997b; Helmke & Weinert, 1997). Intelligenter Schülerinnen und Schüler können sich schneller auf neue Aufgaben einstellen, sie verfügen über effektivere Problemlösestrategien, erkennen leichter lösungsrelevante Regeln, verfügen über eine größere Verarbeitungskapazität und elaboriertere Gedächtnisstrategien (Hasselhorn & Grube, 1997). Auf der Grundlage der SCHOLASTIK-Studie berichten Helmke und Weinert (1997) das in Abbildung 23.6 gezeigte Pfadmodell zum wechselseitigen Zusammenhang zwischen Intelligenz und Schulleistung.

Zwei Befunde erscheinen besonders bemerkenswert.

**(1) Der Einfluss der Intelligenz nimmt ab.** Zum einen zeigt sich, dass es über die Zeit zu einem Absinken der Effekte der Intelligenz auf die Schulleistungen kommt (von  $.30$  auf  $.14$ ) und gleichzeitig die Rolle des Vorwissens wichtiger wird (von  $.45$  auf  $.63$ ). Diese Befundlage ist kongruent zu einer Vielzahl von Studien, in denen die Bedeutung des Vorwissens für schulisches Lernen untersucht wurde (vgl. im Überblick Renkl, 1996). Die prädiktive Bedeutung der Intelligenz ist umso größer, je unbekannter die Lerninhalte sind, das heißt je weniger Vorwissen zu diesen Inhalten bereits vorliegt. Je enger angelehnt aber die Unterrichtsstoffe an das Vorwissen sind, desto unbedeutender ist die Rolle der Intelligenz.

**(2) Schulleistung beeinflusst die Intelligenzentwicklung.** Das zweite sehr interessante Ergebnis in Abbildung 23.6 bezieht sich auf die Pfade

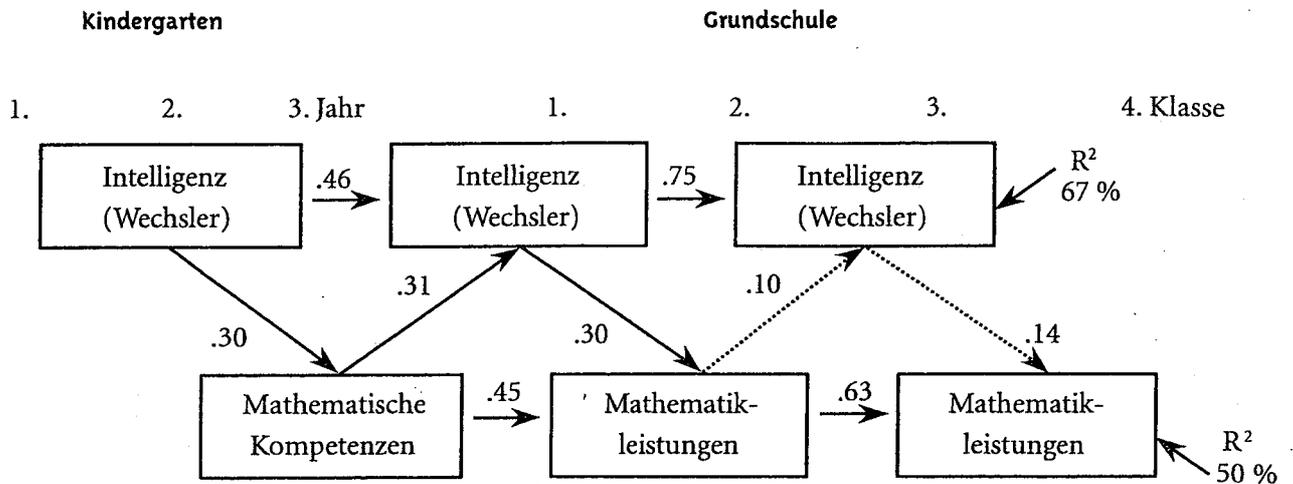


Abbildung 23.6. Längsschnittliches Zusammenhangsgefüge zwischen Schulleistungen und Intelligenz (aus Helmke & Weinert, 1997, S. 108)

von der Schulleistung auf die Intelligenz. Weniger, dass diese schwächer werden, sondern vielmehr, dass sie überhaupt auftreten, ist von besonderem Interesse, implizieren sie doch, dass die schulischen Wissenserwerbsprozesse auch Einfluss auf die Intelligenzentwicklung nehmen. Helmke und Weinert (1997; vgl. auch Ceci, 1991) schlussfolgern auf Grund dieser Befunde, dass die Intelligenz nicht nur eine Bedingung, sondern stets auch eine Folge schulischen Lernens darstellt. Sehr deutliche Belege für diese Schlussfolgerung bieten kürzlich vorgestellte Befunde von Köller und Karim (2001). Die Autoren untersuchten die sprachfreie Intelligenzentwicklung vom Beginn der 7. Jahrgangsstufe bis zum Ende der 10. Jahrgangsstufe in Gymnasien und Realschulen neuer Bundesländer. Die erste Erhebung fand direkt zum Zeitpunkt des Übergangs von der Grundschule in die Sekundarstufe I statt. Neben der sprachfreien Intelligenz in der 7. Jahrgangsstufe und der Schulform wurden auch Indikatoren des sozialen Hintergrunds verwendet, um Intelligenztestwerte am Ende der 10. Jahrgangsstufe vorherzusagen. Mehrebenenanalysen ergaben, dass bei Kontrolle der Ausgangsintelligenz und der familiären Herkunft die Intelligenzentwicklung am Gymnasium deutlich günstiger war, d.h. Gymnasiasten, die in Klasse 7 die gleiche Aus-

gangsintelligenz und soziale Herkunft hatten wie ihre Kameraden von der Realschule; wiesen am Ende der 10. Jahrgangsstufe eine mittlere Intelligenz auf, die 11 IQ-Punkte höher war als auf der Realschule. Mit dem günstigeren Lernklima am Gymnasium gehen offenbar nicht nur bessere Schulleistungsentwicklungen einher, sondern auch eine besondere Ausschöpfung kognitiver Begabungsreserven.

## 5.2 Schulleistungsentwicklung und die Entwicklung von Fähigkeitsselbstkonzepten und schulischen Interessen

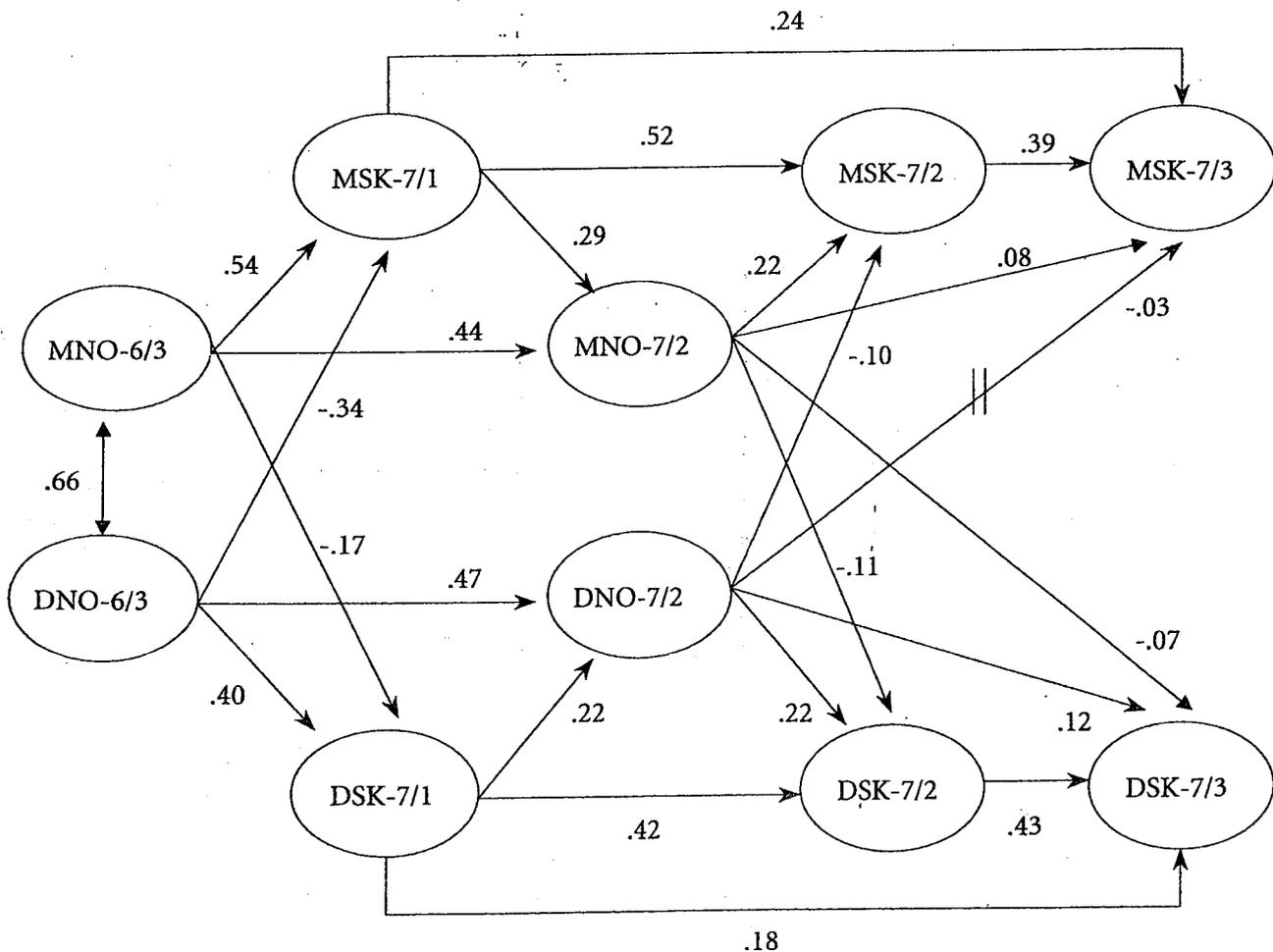
Schulische Selbstkonzepte und Interessen gelten als wichtige Einflussgrößen auf schulisches Lernen, wofür es auch einige Evidenz gibt, wie unten gezeigt wird. Im Vordergrund der nachfolgenden Ausführungen steht aber wiederum das wechselseitige Bedingungsgefüge mit Schulleistungen in einer längsschnittlichen Perspektive. Wir beginnen mit dem Zusammenspiel von Fähigkeitsselbstkonzepten und Schulleistungen.

Mittlerweile liegen hinreichend viele längsschnittliche Arbeiten vor, in denen systematisch die Beziehungen zwischen Schulleistungen und fachspezifischen Selbstkonzepten untersucht wurden (im Überblick Byrne, 1996). Danach ist es

weitgehend unbestritten, dass sich die Entwicklungsverläufe beider Variablen im Kindheits- und Jugendalter gegenseitig beeinflussen (vgl. Helmke & van Aken, 1995; Köller et al., 1999; Marsh, 1990; Marsh & Hattie, 1996; Möller & Köller, 2001).

**Beziehung zwischen Fähigkeit und Selbstkonzept.** Im *skill-development-approach* wird davon ausgegangen, dass sich das Fähigkeitsselbstkonzept in Folge schulischer Leistungserfahrungen und deren Kausalattributionen entwickelt, im *self-enhancement-approach* wird angenommen,

dass das Selbstkonzept vermittelt über motivationale Variablen den Lernprozess fördert und so einen positiven Einfluss auf die Schulleistung nimmt. Als Konsequenz dieser parallel verlaufenden Prozesse kommt es zu der wechselseitigen Beeinflussung, was dann in empirischen Arbeiten zu einem Ergebnismuster führt, wie es in Abbildung 23.7 dargestellt ist. Es handelt sich um Befunde von Köller et al. (1999). Dabei wurden die Fachleistungen (Zeugnisnoten) am Ende der 6. und Mitte der 7. Jahrgangsstufe sowie das akademische Selbstkonzept Anfang, Mitte und Ende



**Anmerkungen:** DNO: Deutschnote; MNO: Mathematiknote; DSK: Akademisches Selbstkonzept in Deutsch; MSK: Akademisches Selbstkonzept in Mathematik; 6/3: Ende Klasse 6; 7/1: Beginn Klasse 7; 7/2: Mitte Klasse 7; 7/3: Ende Klasse 7. N = 1754 Schülerinnen und Schüler verschiedener Schulformen und Bundesländer

**Abbildung 23.7:** Ergebnisse einer Pfadanalyse (voll standardisierte LISREL-8-Lösung) zum längsschnittlichen Zusammenspiel zwischen Fähigkeitsselbstkonzepten und Schulleistungen in Deutsch und Mathematik (aus Köller et al., 1999, S. 131)

der 7. Jahrgangsstufe in Deutsch und Mathematik berücksichtigt. Deutlich wird, dass innerhalb eines Faches die (rekodierten) Leistungsmaße im selben Fach einen positiven Einfluss auf das zeitlich nachgeordnete Fähigkeitsselbstkonzept haben (*skill-development-approach*, z.B. MNO-6/3 auf MSK-7/1). Darüber hinaus zeigen sich im Sinne der *self-enhancement*-Perspektive positive Effekte der zeitlich vorgeordneten Selbstkonzeptvariablen auf die späteren Leistungen, auch wenn frühere Leistungen kontrolliert sind, so dass es sich hierbei also um die Effekte der Fähigkeitsselbstkonzepte auf die Leistungsveränderungen handelt.

Schließlich zeigt die Abbildung 23.7 noch bemerkenswerte *negative* Effekte der Leistungen in dem einen Fach auf das Fähigkeitsselbstkonzept im

anderen Fach, d.h. je höher die Leistungen in dem einen Fach, desto ungünstiger die Entwicklung des Selbstkonzepts in dem anderen Fach. Diesen Prozess hat Marsh (1986) in seinem *Internal/External Frame of Reference Model* beschrieben. Danach werden zwei zentrale Informationsquellen für die Entwicklung fachspezifischer Selbstkonzepte der eigenen Begabung angenommen:

- Der interindividuelle bzw. soziale Vergleich (*external frame of reference*), bei dem die eigenen Leistungen in einem Schulfach mit denen der Mitschülerinnen und Mitschüler verglichen werden (als Folge kommt es zu den positiven Pfaden von der Note auf das Selbstkonzept innerhalb eines Faches).

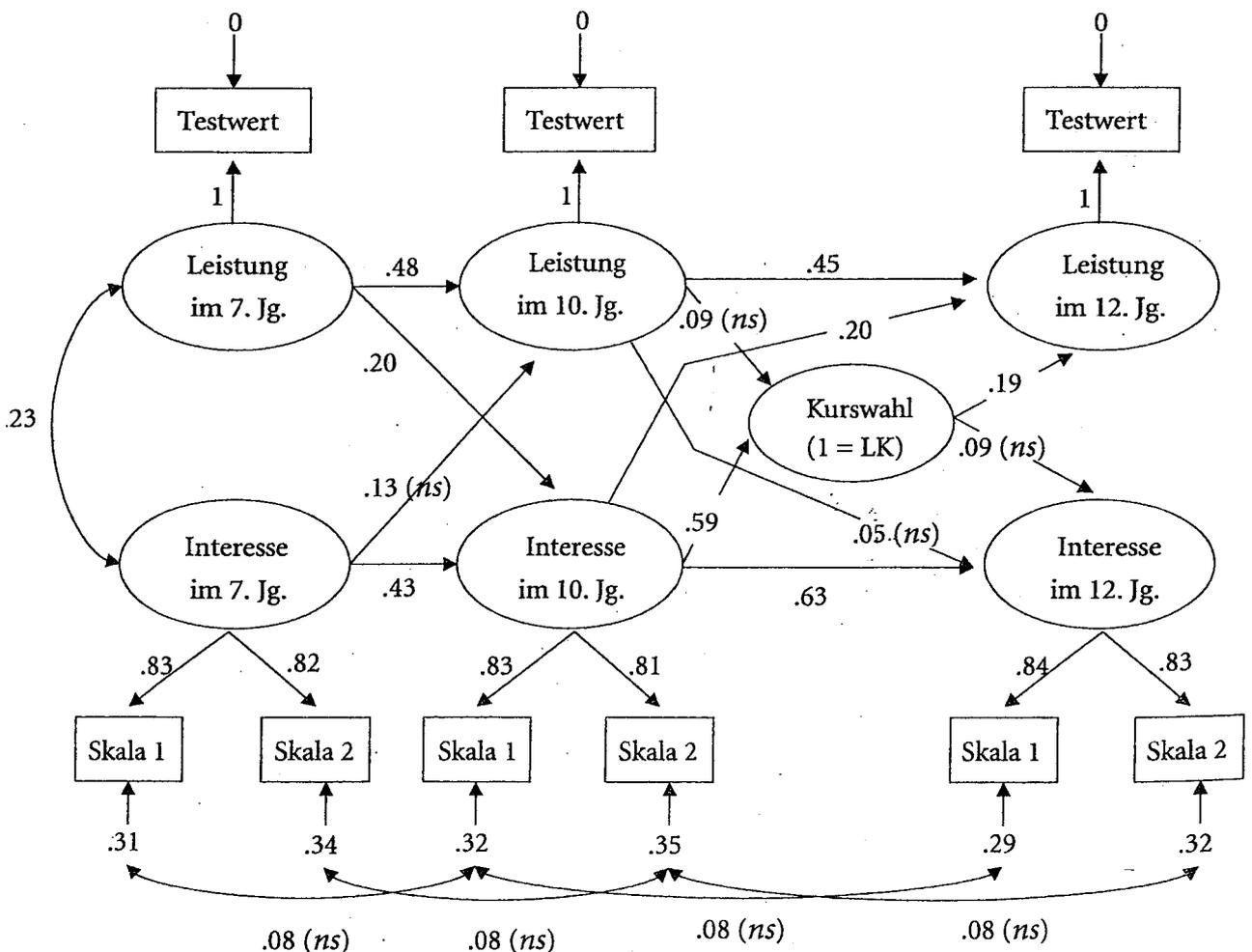


Abbildung 23.8. Strukturgleichungsmodell zum Zusammenhang von Interesse und Testleistung von der 7. bis zur 12. Jahrgangsstufe unter zusätzlicher Berücksichtigung des Kursniveaus (vgl. Köller et al., 2000, S. 176)

- Der intraindividuelle oder auch dimensionale Vergleich (*internal frame of reference*), bei dem Schüler die Leistungsergebnisse in einem Fach mit ihren eigenen Leistungen in einem anderen Fach vergleichen (als Folge kommt es zu den negativen Pfaden von der Note in einem Fach auf das Selbstkonzept im anderen Fach).

### Beziehung von Schulleistung und Interesse.

Auch für die wechselseitige Beeinflussung von Schulleistungen und Interessen liegen mittlerweile längsschnittliche Befunde vor (Köller, Baumert & Schnabel, 2000; Köller et al., 2001). Interesse wird dabei als Form der fachspezifischen intrinsischen Motivation verstanden, die Lernprozesse in der Schule initiieren und aufrechterhalten kann (vgl. Krapp, 1998; Schiefele, 1996; Schiefele & Wild, 2000). Viele Arbeiten belegen aber auch, dass Interessen stark durch wahrgenommenes schulisches Kompetenzerleben beeinflusst werden können (z.B. Baumert, Schnabel & Lehrke, 1998). Die Abbildung 23.8 zeigt über einen Zeitraum von 5 Jahren das Zusammenspiel von Fachleistungen und Interessen in Mathematik. Die Befunde stammen aus einer Arbeit von Köller et al. (2000) und basieren auf einer Stichprobe von über 600 Gymnasiasten.

Erkennbar ist das reziproke Zusammenhängegefüge, das heißt beide Variablen beeinflussen sich gegenseitig über die Zeit, wiewohl einige der zeitversetzten Pfade das Signifikanzniveau von 5 Prozent verfehlen. Zusätzlich zeigt sich noch ein signifikanter Effekt vom Interesse am Ende der Sekundarstufe I auf die Chance, einen Leistungskurs zu wählen.

#### FAZIT

Untersuchungen zur Schulleistungsentwicklung greifen zu kurz, wenn sie zusätzliche Personmerkmale ignorieren, deren Entwicklung unmittelbar mit der Entwicklung fachbezogenen Wissens verbunden ist.

## 6 Methodische Probleme bei der Untersuchung von Schulleistungen und deren Entwicklung

Dieser abschließende Abschnitt hat die Aufgabe, auf methodische Besonderheiten bei der Erforschung von schulischen Bildungsprozessen hinzuweisen, die üblicherweise in anderen Bereichen der entwicklungspsychologischen Forschung in dieser Form nicht zu berücksichtigen sind. Ein Teil dieser Probleme rührt wiederum daher, dass es sich bei Schulleistungen um keine klar umrissenen psychologischen Konstrukte handelt. Bei den weiteren Ausführungen verzichten wir allerdings auf die Darstellung elementarer Gütekriterien von Schulleistungstests wie Reliabilität, Konstruktvalidität und Objektivität (vgl. hierzu Heller & Hany, 2001).

### 6.1 Zur curricularen Validität von Schulleistungstests

Wie bereits in Abschnitt 2 argumentiert wurde, sind es weniger die psychologischen Theorien als vielmehr die vorgegebenen Curricula, die bestimmen, welche Inhalte in Schulleistungstests thematisiert werden können. Im Rahmen internationaler Schulleistungsuntersuchungen ist man dazu übergegangen, das Curriculum eines Landes vierstufig als *intendiertes Curriculum*, *potentielles Curriculum*, *implementiertes Curriculum* und *erreichtes Curriculum* darzustellen (vgl. Baumert, Köller et al., 2000; Baumert, Lehmann et al., 1997; Travers & Westbury, 1989). In der Bundesrepublik Deutschland ist das intendierte Curriculum durch die Lehrpläne und Prüfungsvorschriften definiert. Die zugelassenen Lehrbücher repräsentieren das potentielle Curriculum. Als implementiertes Curriculum gilt der in einer Schule tatsächlich behandelte Stoff. Das erreichte Curriculum schließlich wird durch die Schülerleistungen selbst angezeigt. Ein Schüler bzw. eine Schülerin kann – sofern Lernen der curricular vorgegebenen Stoffe primär

in der Schule stattfindet – nur die Inhalte gelernt haben, die auch behandelt wurden, sich also auf das implementierte Curriculum beziehen. Für die Erforschung von Schulleistungsentwicklungen tritt hier natürlich das Dilemma auf, dass nicht spezielle Leistungstests für das implementierte Curriculum in jeder einzelnen Klasse jeder Schulform in jedem Bundesland konstruiert werden können. Vielmehr müssen sich die Tests am intendierten Curriculum in der Hoffnung orientieren, dass eine hohe Übereinstimmung mit dem implementierten Curriculum besteht. Systematische Analysen zur curricularen Validität von Aufgaben in Schulleistungstests, die im Rahmen der dritten internationalen Mathematik- und Naturwissenschaftsstudie (TIMSS) durchgeführt wurden, weisen auf eine erhebliche Übereinstimmung mit dem realisierten Curriculum hin, so dass wenigstens in neueren Untersuchungen davon ausgegangen werden kann, dass die gemessenen Merkmale eine zufriedenstellende curriculare Validität (bezogen auf das intendierte und implementierte Curriculum) aufweisen (vgl. Baumert et al., 2000a, 2000b; Baumert, Lehmann et al., 1997).

## 6.2 Probleme bei der statistischen Modellierung von schulischen Entwicklungsverläufen

Lernen im Sinne der relativ dauerhaften Veränderung von kognitiven Strukturen (s.o.) meint bezogen auf Schulleistungen, dass es durch systematisch hergestellte Lehr-Lernsituationen auf Seiten der Schülerinnen und Schüler zum Aufbau, zur Veränderung und/oder Verfestigung von fachspezifischen Wissensstrukturen kommt. Greeno (1989) spricht von der *Reorganisation* von Wissensstrukturen und Kintsch (1986, 1988) konzeptualisiert Lernen als die Umstrukturierung von mentalen Situationsmodellen (vgl. im Überblick Steiner, 1997). Bei allen diesen Konzeptionen werden *höhere Lernprozesse* (vgl. Resnick, 1987) thematisiert, die nicht mit Reiz-Reaktions-Mustern erklärt werden können; „sie sind zumeist von

komplexer Art, beruhen auf einer differenzierten Beurteilung der Lernsituation, implizieren oft eher heuristische als algorithmische Verfahren und sind nicht selten durch eine Vielfalt von Lösungswegen und Lösungsmöglichkeiten gekennzeichnet, manchmal auch durch eine beachtliche Unsicherheit in bezug auf den Umgang mit den zu lernenden Inhalten“ (Steiner, 1997, S. 279).

**Lernen als Prozess.** Lernen muss in diesem Zusammenhang als ein Prozess verstanden werden, in dem die verschiedenen Komponenten *Verstehen, Speichern, Abrufen* und *Anwenden* verknüpft werden (Danserau, 1985). Bezogen auf die Messung und Analyse der schulischen Leistungsentwicklung impliziert dies, dass man Tests und Untersuchungsdesigns wählen muss, die eine Diagnose der Veränderung der curricular definierten Wissensstrukturen auf Seiten der Schülerinnen und Schüler ermöglichen. Beherrscht ein Schüler zu einem Zeitpunkt  $T_0$  die Lösung einer Aufgabe nicht, wohl aber zu einem späteren Zeitpunkt  $T_1$ , so kann davon ausgegangen werden, dass in der Zwischenzeit ein schulischer Lernprozess stattgefunden hat, sofern nicht die Annahme besteht, die erfolgreiche Lösung der Aufgabe sei die bloße Folge eines kognitiven Reifungsprozesses. Schulische Bildungsprozesse können dementsprechend nur mit Hilfe von Längsschnittstudien untersucht werden, in denen die *Veränderung des Wissens* als Folge des Lehr-Lerngeschehens in der Schule die *zentrale Untersuchungsvariable* ist. Wählt man dabei eine langfristige Untersuchungsperspektive, beispielsweise den Zeitraum der gesamten Sekundarstufe I und II (also von der 5. bis einschließlich der 13. Jahrgangsstufe), so wird unmittelbar deutlich, dass in der 13. Jahrgangsstufe nicht mehr dieselben Testaufgaben verwendet werden können wie in der 5. Jahrgangsstufe oder gar umgekehrt. Die Vorgabe einer Aufgabe zur Differentialrechnung in der 5. Jahrgangsstufe wird wegen der zu erwartenden Lösungswahrscheinlichkeit von  $p = 0$  wenig diagnostische Information liefern. Das Gleiche gilt für die Bearbeitung einer einfachen Aufgabe zur Flächenberechnung eines Quadrats in der 13. Jahrgangsstufe, die dort

### 5. Jahrgangsstufe

Aufgaben 1–15	Aufgaben 16–30
------------------	-------------------

### 7. Jahrgangsstufe

Aufgaben 16–30	Aufgaben 31–45
-------------------	-------------------

### 9. Jahrgangsstufe

Aufgaben 31–45	Aufgaben 46–60
-------------------	-------------------

Abbildung 23.9.

Schematische Darstellung eines Ankeritem-Designs zur Erfassung der Schulleistungsentwicklung von der 5. bis zur 9. Jahrgangsstufe

vermutlich mit einer Wahrscheinlichkeit von  $p = 1$  gelöst wird.

**Ankeritem-Designs.** Den Ausweg aus diesem methodischen Dilemma bieten Ankeritem-Designs, d.h. Teilmengen der Aufgaben (Ankeritems) werden zu benachbarten Messzeitpunkten wiederholt vorgegeben und durch zusätzlich eingesetzte Aufgaben ergänzt. Die Abbildung 23.9 zeigt solch ein Design für den Fall mit drei Messzeitpunkten in der Sekundarstufe I.

Die Aufgaben 16 bis 30 in Abbildung 23.9 werden zum ersten und zweiten Messzeitpunkt vorgegeben, die Items 31 bis 45 zu den letzten beiden. *Item-Response-Modelle* (Hambleton & Swaminathan, 1989; Rost, 1996) erlauben hier, auf der Basis der Ankeritem-Informationen Transformationsvorschriften zu bestimmen, mit deren Hilfe man die Schülerleistungen der verschiedenen Erhebungszeitpunkte auf einem gemeinsamen Maßstab abtragen kann. Auf diese Weise wird es möglich, intraindividuelle Entwicklungsverläufe über längere Zeiträume zum Gegenstand der Analysen zu machen. Diese Form der systematischen Analyse von Veränderungen im Schulleistungsbereich ist allerdings voraussetzungsvoll (vgl. Baumert, Lehmann et al., 1997):

- (1) Die in den Ankeritems thematisierten Stoffe sollten zwischen den beiden Messzeitpunkten unterrichtet werden. Gleichzeitig sollte es aber schon einer kleineren Schülergruppe möglich sein, die Items vorher zu lösen.
- (2) Sofern die in den Ankeritems thematisierten Stoffgebiete schon zu früheren Zeitpunkten durchgenommen wurden, sollten sie in den untersuchten Schuljahren bei der Erarbeitung des neuen Stoffs wieder aufgenommen, kon-

solidiert und besser integriert werden. Fehlt diese inhaltliche Kohärenz über die Jahrgangsstufen, so ist die Diagnose von Wissenszuwächsen wenig sinnvoll.

Trotz dieser methodischen Schwierigkeiten ist man auf derartige Erhebungsdesigns angewiesen, will man sich individuellen Bildungsverläufen aus einer empirischen Perspektive nähern.

## 6.3 Die Rolle des Antwortformats in Schulleistungstudien

Im schulischen Alltag bearbeiten Schülerinnen und Schüler offene Aufgaben und Probleme. Lösungen müssen nicht nur aufgeschrieben werden, es kommt oftmals ein Erläuterungsprozess hinzu, in dem beispielsweise der Lösungsprozess nachgezeichnet werden muss. Standardisierte, curricular validierte Tests zur Schulleistungsdia- gnose sehen dagegen oftmals ein Antwortformat mit Mehrfachwahlen – *multiple choice* (MC) – vor, bei dem ein Schüler aus 2 bis 5 Antwortalternativen die richtige Lösung einer Aufgabe herausfinden und ankreuzen muss (zur Vielfältigkeit von Antwortformaten s. Klauer, 2001).

**Vorteile und Probleme von MC-Aufgaben.** Der Vorteil dieses Antwortformats liegt in der hohen Auswertungsökonomie und Objektivität. Häufig ist kritisiert worden (s. im Überblick Klauer, 2001), dass MC-Aufgaben für Schülerinnen und Schüler ungewohnt seien, da Haus- und Klassenarbeitsaufgaben immer ein offenes Format hätten. Dadurch würde letztendlich die ökologische Validität der Befunde reduziert. Aus wissenschaftlicher Sicht (Helmke, 1997a) ist über lange Zeit

auch eingewandt worden, dass Mehrfachwahlaufgaben im wesentlichen nur reproduktive Leistungen, nicht aber produktive Fähigkeiten – schlussfolgerndes Denken, komplexe Operationen oder gar Problemlösen – erfassen können. Vermisst wurde bei MC-Aufgaben auch die Möglichkeit, systematische Fehleranalysen durchzuführen (Helmke, 1997a). Als bessere Alternative gelten offene Antwortformate.

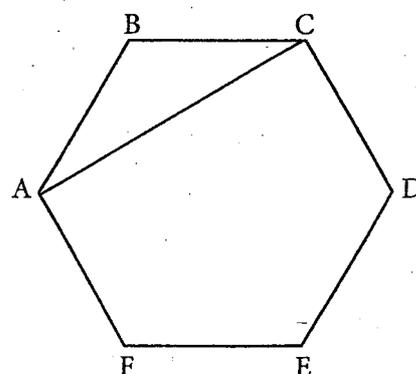
**Das Antwortformat ist von nachgeordneter Bedeutung.** Ein Blick in die jüngere methodologische Literatur (z.B. Frary, 1985; Thissen et al., 1994) zeigt allerdings, dass das Antwortformat eher eine nachgeordnete Bedeutung hat. So fassen Thissen et al. (1994, S. 113) ihre Ergebnisse zusammen: "Restricted factor analysis shows that the free-response sections measure the same underlying proficiency as the multiple-choice sections". Auch muss festgehalten werden, dass MC-Aufgaben sehr gut geeignet sind, um mit Hilfe entsprechend formulierter falscher Antwortalternativen so genannte Misskonzeptionen von Schülern zu diagnostizieren. Bei Misskonzeptionen handelt es sich um Alltagsvorstellungen von Schülerinnen und Schülern, die mit zu erwerbenden Unterrichtsinhalten inkompatibel sind und daher den Lernprozess oftmals hemmen (vgl. im Überblick Perkins & Simmons, 1988). Hierzu zählen beispielsweise die Annahmen, die Erde sei eine Scheibe oder die Sonne würde sich um die Erde drehen.

**Die Effekte des Antwortformats müssen psychologisch erklärt werden.** Im übrigen bleiben potentielle Effekte des Antwortformats inhaltlich unbestimmt, solange keine eindeutigen Hypothesen über psychologische Prozesse vorliegen, die das Antwortverhalten von Schülern erklären. Im Falle von MC-Items kann man allerdings relativ einfach spezifische Annahmen über systematische Effekte des Antwortformats machen. So wird das Format bei diesen Items immer dann für den Lösungsprozess wichtig, wenn eine Person die Antwort auf eine Testfrage nicht weiß. Strategien, die in diesem Fall verwendet werden können, sind zum einen Raten, d.h. zufälliges Wählen einer

Antwortalternative (Köller, 1994), zum anderen intelligentes Ausschließen von Antwortalternativen, indem aus allen Antwortalternativen diejenige mit subjektiv höchster Plausibilität ausgewählt wird (Lord, 1974). Dies kann die Aufgaben dann in der Tat hinsichtlich ihrer Schwierigkeit und Konstruktvalidität erheblich verändern. Lord (1974) weist darauf hin, dass besonders attraktive Distraktoren unter den Antwortalternativen gern gewählt werden, was bei sehr schweren Items oftmals dazu führt, dass der Prozentsatz richtiger Lösungen unter die Ratewahrscheinlichkeit fällt. Eine schönes Beispiel, das zeigt, wie durch das MC-Format Items leichter werden können, indem sich die zur Lösung erforderlichen kognitiven Operationen verändern, bietet die in Abbildung 23.10 gezeigte Aufgabe zur Elementargeometrie aus der TIMSS-Studie (vgl. Klieme, 2000, S. 92).

Zur formalen Lösung dieser Aufgabe wären Kenntnisse der Begriffe „reguläres Sechseck“ und „Diagonale“ sowie die Anwendung eines Satzes aus der Satzgruppe des Pythagoras nötig. Tatsächlich reicht es jedoch, unter Berücksichtigung der vorgegebenen Antwortalternativen, sich zu überlegen, dass die Strecke von AC wohl länger ist als eine Seite des Sechsecks (also  $> 10$  cm), aber kürzer als die Summe zweier Seitenlängen

Alle Seiten des regulären Sechsecks ABCDEF sind 10 cm lang. Wie groß ist die Länge der Diagonale AC?



- A:  $10\sqrt{3}$  cm
- B: 20 cm
- C:  $5\sqrt{3}$  cm
- D: 10 cm
- E:  $20\sqrt{3}$  cm

**Abbildung 23.10.** Multiple-choice-Aufgabe aus dem TIMSS-Test zur voruniversitären Mathematik (vgl. Klieme, 2000, S. 92)

(also  $< 20$  cm). Damit bleibt nur die Antwortalternative A als richtige Lösung übrig ( $10 \cdot \sqrt{3}$  cm = 17,3 cm). Bei einem offenen Antwortformat wäre die Anwendung des Satzes von Pythagoras unumgänglich und die Aufgabe damit erheblich schwerer. Untersuchungen zu Schulleistungen bzw. deren Entwicklung lösen zum Teil solche Probleme, indem beide Antwortformate (*multiple-choice* und offene Antworten) verwendet werden.

#### 6.4 Zur Dimensionalität von Schulleistungen bzw. Schulleistungstests

Die Frage nach der Dimensionalität von fachspezifischen Schulleistungstests ist aus einer psychologischen wie fachdidaktischen Perspektive interessant. Psychologisch stellt sich für ganz unterschiedliche Teilbereiche und Anforderungen eines Schulfaches die Frage, ob es eine gemeinsame Dimension ist, die die Leistungen in diesen verschiedenen Bereichen determiniert oder ob es differenzielle intellektuellen Faktoren sind, die das Lösungsverhalten steuern. Aus didaktischer Perspektive wird nicht selten argumentiert, dass die Stoffgebiete (z.B. Leseverständnis, Wortschatz, Grammatik in sprachlichen Fächern; Geometrie, Algebra, Wahrscheinlichkeitsrechnung in Mathematik) so grundlegend verschiedene Inhalte haben, dass man sie nicht „über einen Kamm scheren“ dürfe. Frühe Beiträge zu dieser Dimensionalitätsproblematik stammen von Roeder und Treumann für den Mathematik- und Fremdsprachenunterricht (1974a, b). Quintessenz der Reanalysen und Literatur-Reviews der beiden Autoren ist, dass die Leistungen in beiden Bereichen mehrdimensional sind. So ergaben z.B. Reanalysen einer Studie von Otter (1968) für den Fremdsprachenunterricht am Ende der Sekundarstufe I in England drei Dimensionen (vgl. Roeder & Treumann, 1974a, S. 66f): (1) Rezeptives Sprachverständnis, (2) Lesefertigkeiten und (3) schriftlicher Ausdruck. In Mathematik ergaben sämtliche gesammelten Studien keine ein-

dimensionalen Lösungen, bemerkenswert ist allerdings, dass alle Untersuchungen, in denen die Items eng am Curriculum orientiert waren, einen starken Generalfaktor in den entsprechenden Faktorenanalysen aufwiesen.

**Drei-Faktoren-Modell oder g-Faktor?** Sang et al. (1986) analysierten die Englischleistungen von rund 14000 Gymnasiasten. Konfirmatorische Faktorenanalysen ergaben Evidenz für ein Drei-Faktorenmodell mit den Dimensionen Elementares Wissen (Aussprache, Orthographie, Wortschatz), komplexeres Wissen (Grammatik, Leseverständnis) und Kommunikationsfertigkeiten (Hörverständnis, Sprachproduktion). Die Interkorrelationen dieser drei Faktoren (zwischen .60 und .84) belegen aber auch einen hohen gemeinsamen Varianzanteil, so dass vermutet werden kann, dass ein Modell mit einem starken g-Faktor und drei spezifischen (unkorrelierten) Faktoren ebenso gut die empirischen Zusammenhänge erklärt hätte.

**Evidenz für einen g-Faktor.** In Auswertungen der TIMSS-Mittelstufendaten (vgl. Köller, 1998) wurden Leistungen aus sechs unterschiedlichen mathematischen Sachgebieten (Zahlen und Zahlenverständnis, Geometrie, Algebra, Darstellung und Analyse von Daten/Wahrscheinlichkeitsrechnung, Messen und Maßeinheiten, Proportionalität) schulformenspezifisch faktorenanalysiert. In allen Schulformen der Sekundarstufe I ergab sich erheblich Evidenz für einen g-Faktor allgemeiner mathematischer Kompetenz. Ähnliche Befunde berichten Köller (1998) für Englisch und Klieme (2000) für voruniversitäre Mathematik- und Physikleistungen. Diese Befundlage impliziert insgesamt, dass die fachspezifischen Leistungen in Schulleistungstests durch ein komplexeres Fähigkeitssyndrom beeinflusst werden, so dass in empirischen Untersuchungen innerhalb eines einzelnen Faches curriculare oder kognitive Teilkomponenten analytisch kaum trennbar sind, man bei der Untersuchung von schulischen Leistungsverläufen also keinen substantiellen Fehler macht, wenn man die Leistungen innerhalb eines Faches als unidimensional ansieht.

## FAZIT

Bei der Beschäftigung mit Schulleistungen im Rahmen psychologischer Fragestellungen treten sowohl inhaltliche als auch methodische Besonderheiten bzw. Probleme auf, die nicht leicht lösbar sind und die Interpretierbarkeit empirischer Befunde erschweren.

## 7 Zusammenfassung

Dieser Beitrag hatte die Beschreibung der Entwicklung von Schulleistungen zum Gegenstand. Dabei sollte deutlich geworden sein, dass Schulleistungen keine theoretisch so „griffigen“ Konstrukte sind, wie es beispielsweise die psychometrische Intelligenz ist. Vielmehr sind es die Curricula, die das Konstrukt Schulleistung zu einem erheblichen Teil definieren. Die Bestimmung der Entwicklungsverläufe und die Definition von Kompetenzniveaus oder -stufen aus einer psychologischen Theorie heraus sind somit schwierig und gelingen eher durch entsprechende *post-hoc*-Analysen, wie sie hier in Form des *scale anchorings* vorgestellt wurden.

Erkennbar sollte in diesem Kapitel auch geworden sein, dass schulische Entwicklungsverläufe sehr stark an systematisch hergestellte Lerngelegenheiten in Schulen gebunden sind, ohne deren

Verfügbarkeit ein breiter Kranz von Kompetenzen kaum erworben werden kann. Dieser Punkt wird aus unserer Sicht viel zu selten betont, wenn es um Antezedenzen von Schulleistungen geht.

Schließlich wurde gezeigt, dass Schulen als Entwicklungsmilieus keineswegs nur Einfluss auf die Entwicklung von Schulleistungen nehmen, sondern vielmehr mit einem breiten Wirkungsgrad auf vielfältige Personmerkmale zu rechnen ist.

### Weiterführende Literatur

- Deutsches PISA-Konsortium (Hrsg.) (2001). PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske+Budrich. Das Buch gibt sehr detailliert Auskunft über die Erfassung und den Stand schulischer Leistungen in den Bereichen Lesen, Mathematik und Naturwissenschaften am Ende der Sekundarstufe I. Dabei wird auch ausführlich auf die Definition der Kompetenzstufen in den drei Fächern eingegangen.
- Scheerens, J., & Bosker, R. (1997). The foundations of educational effectiveness. Oxford: Elsevier. Dieses Buch beschreibt eine Vielzahl von Arbeiten, in denen der Einfluss von Schule und Unterricht nicht nur auf die Leistungsentwicklung sondern auch auf die Entwicklung motivationaler und psychosozialer Variablen untersucht wurde.
- De Corte, E. & Weinert, F. E. (Hrsg.) (1996). International Encyclopedia of Developmental and Instructional Psychology. In dieser Enzyklopädie wird neben allgemeinen entwicklungspsychologischen Ansätzen auch speziell auf Aspekte der Entwicklung in schulischen Kontexten eingegangen.

# Frühe Differenzierung fördert leistungsstarke Schüler

Jetzt haben wir es schriftlich! Noch dazu von jenen Leuten, die es am besten wissen. Und es wurde an maßgebender Stelle veröffentlicht: »Frühe Differenzierung fördert leistungsstarke Schüler.« Die neueste, vollständig überarbeitete Ausgabe des pädagogischen Standardwerkes

»Entwicklungspsychologie« von Oerter und Montada enthält einen Beitrag der Professoren Köller und Baumert zum Thema »Entwicklung schulischer Leistungen« (S. 756-786), mit eben dieser Feststellung. Sie hat bisher noch nicht die gebührende Beachtung gefunden.

von Ulrich Sprenger

**K**öller und Baumert referieren zunächst, dass in anderen Ländern ein gegliedertes Schulsystem mit einer Gruppierung der Schüler nach Fähigkeiten keine Vorteile bringt. Dann aber folgt die überraschende Mitteilung (S. 770), das sei zumindest in Deutschland anders: *»Implizieren diese internationalen Befunde, dass die Fähigkeitsgruppierungen wenigstens im Hinblick auf optimale Leistungsförderungen ineffektiv sind, so ergibt sich in deutschen Studien ein abweichendes Bild.«*

Belegt wird diese Feststellung mit Daten aus der ersten großen Studie des 1963 gegründeten Max-Planck-Institutes für Bildungsforschung (MPIB). Es handelt sich um das Projekt »Schulleistung«, auch »Gymnasiastenstudie« genannt. Für diese sehr aufwendige Studie sind 1968-1970 in den zehn alten Ländern der Bundesrepublik und in West-Berlin etwa 14.000 Gymnasiasten an etwa 450 Gymnasien je drei Tage zu Beginn und am Ende des siebten Jahrgangs untersucht worden, und zwar im Hinblick auf ihre Leistungen in den Fächern Deutsch, Englisch, Mathematik, außerdem auch noch im Hinblick auf ihre intellektuellen Grundfähigkeiten, auf ihr Freizeitverhalten sowie im Hinblick auf den Bildungsstand der El-

tern- und der Großeltern-Generation.

Eine Teil-Auswertung dieser Daten war von Professor Peter M. Roeder, dem Vorgänger Baumerts als Direktor am MPIB, schon 1997 in dem von Weinert und Helmke herausgegebenen Standardwerk »Entwicklung im Jugendalter« veröffentlicht worden. Auf diese Veröffentlichung und auf eine frühere, 1991 von Roeder und Sang herausgegebene Auswertung berufen sich Köller und Baumert nun bei ihrem Beweisgang (S. 770): *»Ausgangspunkt der Arbeit war die Fragestellung, ob eine spätere Differenzierung (nach der sechsten Jahrgangsstufe) bei leistungsstarken Schülern im Vergleich zur Differenzierung nach vier Schuljahren ungünstige Effekte auf die Schulleistung hat. Dazu wurden die Leistungen von Gymnasiasten in der siebten Jahrgangsstufe in Deutsch, Mathematik und Englisch aus einem Bundesland mit sechsjähriger Grundschule mit denen der übrigen Bundesländer mit vierjähriger Grundschule verglichen. Die Tabelle 23.2 zeigt die Resultate der Studie.«* (Das Land mit sechsjähriger Grundschule war West-Berlin.)

Die Tabelle 23.2 wird im Text nicht kommentiert. Daher ist nur für jemanden, der mit dieser Art Statistiken einigermaßen vertraut ist, zu erkennen: **In Englisch und Mathematik liegen die Leistungsunter-**

**schiede bei etwa einer Standardabweichung. Das ist – wie auf Anfrage bestätigt wurde – ein Wissensvorsprung von etwa anderthalb Schuljahren!**

Die Veröffentlichung von 1991 enthält auch Angaben über die weitere Entwicklung der beiden Schülergruppen und den am Ende des siebten Jahrgangs vorgefundenen Leistungsstand: *»Im Englischen liegt der Unterschied immer noch bei etwa einer Standardabweichung, bleibt also fast unverändert bestehen. Im Mathematik- und Deutschunterricht geht die Differenz auf etwa eine dritte Standardabweichung zurück, ist also immer noch auch praktisch bedeutsam. (...) Die Aussagen der Lehrer über ihre Unterrichtsplanung für das siebte Schuljahr und über die Themen und Aufgaben, die sie späteren Schuljahren zuweisen, deuten darauf hin, dass der Leistungsausgleich zwischen beiden Gruppen langsamer vonstatten gehen wird, als man bei gleich befähigten Schülern vermuten würde.«* (Roeder und Sang 1991, S. 167/68)

Vor diesem Hintergrund bekommt die folgende, von Köller und Baumert vorgelegte Zusammenfassung (S. 770) von Ergebnissen des Projektes »Schulleistung« ein erheblich höheres Gewicht: **»Unübersehbar zeigen sich deutliche Leistungsvorteile der Gymnasiasten aus den Bundesländern mit vierjähriger**

*Grundschule. Die frühere Differenzierung scheint also leistungsstärkere Schüler in allen drei Fächern besser zu fördern.«*

Offenbar sind diese Ergebnisse nach wie vor gültig. Denn die beiden Autoren halten es nicht für erforderlich, sie mit neueren Daten zum niedrigeren Fördererfolg von undifferenzierten Lerngruppen zu bestätigen, zum Beispiel mit den entsprechenden Daten des 1991 gestarteten Projektes BIJU (»Bildungsprozesse und psychosoziale Entwicklung im Jugendalter und jungen Erwachsenenalter«). Stattdessen wird durch einen Vergleich von Gymnasiasten mit gleich begabten Realschülern erklärt, warum leistungsstarke Schüler in Deutschland an Gymnasien generell eine bessere Förderung erfahren. Dieser Vergleich ist Teil einer im Jahre 2001 verfassten »methodisch elaborierteren Studie zu Effekten der Leistungsgruppierung im deutschen Schulsystem« von Köller und Baumert. Er wurde durchgeführt anhand von Daten eben jenes Projektes BIJU.

Ergebnis dieser Studie aus dem Jahre 2001: Gymnasiasten haben gegenüber gleich begabten Realschülern, mit denen sie am Anfang des siebten Jahrgangs noch denselben Leistungsstand hatten, am Ende des zehnten Jahrgangs zum Beispiel in Mathematik einen Wissensvorsprung von »mehr als einer halben Standardabweichung«. (Das ist fast ein ganzes Schuljahr!)

Es folgt (S. 771) eine weitere, ebenfalls recht überraschende Auswertung: *»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 Gymnasialklassen, als vielmehr Folge der schulformspezifischen Lehrerbildung im deutschen Lehrerbildungssystem. – Zusammenfassend zeigt sich zumindest für das deutsche Schulsystem, dass bezogen auf die Fachleistungsentwicklung leistungsstarke Schüler von der Differenzierung im Sekundarbereich profitieren.« (»makroadaptiv« ist eine Ausrichtung des Lehrerverhaltens am durchschnittlichen Leistungsvermögen der gesamten Klasse, »mikroadaptiv« ist die Ausrichtung am Leistungsvermögen der einzelnen Schüler, wie dies bei Binnendifferenzierung geschieht.)

Im letzten Teil des Beitrags (S. 778) gibt es eine dritte, ebenfalls überraschende Information (Vorweg-Kommentar: »...bei Kontrolle der Ausgangsintelligenz und der familiären Herkunft« heißt: bei ähnlicher Intelligenz zu Beginn des siebten Jahrgangs und bei einer ähnlichen familiären Herkunft.):

»Mehrebenenanalysen ergaben, dass bei Kontrolle der Ausgangsintelligenz und der familiären Herkunft die Intelligenzentwicklung am Gymnasium deutlich günstiger war, d.h. Gymnasiasten, die in Klasse sieben die gleiche Ausgangsintelligenz und soziale Herkunft hatten wie ihre Kameraden von der Realschule, wiesen am Ende der zehnten Jahrgangsstufe eine mittlere Intelligenz auf, die 11 IQ-Punkte höher war als auf der Realschule. Mit dem günstigen Lehrklima am Gymnasium gehen offenbar nicht nur bessere Schulleistungsentwicklungen einher, sondern auch eine besondere Ausschöpfung kognitiver Begabungsreserven.«

Die von Köller und Baumert vorgelegten Erkenntnisse des MPIB und deren Veröffentlichung in einem Standardwerk der Pädagogik erfordern unseres Erachtens folgende Konsequenzen:

1. Es sollte ab jetzt niemand mehr von der Einführung sechs-

jähriger Grundschulen eine Verbesserung der Lernleistungen deutscher Schüler erwarten – und Entsprechendes ausrufen dürfen.

2. Eine über den vierten Jahrgang hinausgehende Beschulung leistungsstarker Schüler in undifferenzierten, leistungsgemischten Lerngruppen ist nach Bekanntwerden dieser MPIB-Befunde als eine vorsätzliche Benachteiligung leistungsstarker Schüler zu betrachten – und daher einzustellen.

3. Ähnliches gilt für die Unterrichtung leistungsschwächerer Schüler in undifferenzierten, leistungsgemischten Lerngruppen. Köller und Baumert verweisen (S. 772) auf die Arbeiten von Schwarzer und Jerusalem aus dem Jahre 1982, denen zufolge auch die leistungsschwächeren Schüler von einer früheren Differenzierung profitieren, weil sie in den leistungshomogeneren Gruppen von Hauptschulen ein höheres Selbstwertgefühl entwickeln können.

4. Ebenso wie Gymnasiasten und Hauptschüler dürften auch Realschüler als leistungsstärkere Schüler von einer frühen Differenzierung und von einer schulformbezogenen Ausbildung ihrer Lehrer profitieren. Das stellt die in Regionalschulen, Regelschulen, Mittelschulen und Haupt-, Realschulen übliche, erst nach dem sechsten Jahrgang einsetzende Differenzierung in Frage, ebenso auch die Tatsache, dass es in NRW eine schulformbezogene Ausbildung der Realschullehrer nicht mehr gibt.

5. Die Frage, ob die Gemeinnützige Gesellschaft Gesamtschule e.V. so gemeinnützig ist, wie sie zu sein vorgibt, steht ab jetzt im Raum. Sie hat mit ihren Aktivitäten – wie durch die MPIB-Studien bewiesen wurde – folgenreiche Fehlinvestitionen veranlasst und seit drei Jahrzehnten eine Vernachlässigung leistungsstärkerer Schüler betrieben – und will damit nicht aufhören.



Gruppen- und Studienreisen GmbH  
 CTS - Herforder Straße 75 - 32657 Lemgo  
 Tel. 05261/2506-0 - Fax 05261/16300  
 E-mail: info@cts-reisen.de  
 Internet: www.cts-reisen.de

**Vive la France!**  
**Klassenfahrt nach Paris**  
**5 Tage ab € 145**  
[www.cts-sparpreise.de](http://www.cts-sparpreise.de)

**Die Multimediale Lernumgebung für Ihre Schule.**

Planung | Realisierung | Wartung & Betreuung

[www.mts-reinhardt.de](http://www.mts-reinhardt.de)



IT-Komplettlösung für jede Schule

**KURHAUS DR. LAUTER**  
 Gesundheit erleben ... BAD KOHLGRUB ... voll behindertfähig!

herzliches Ambiente, physikalische Therapie,  
 MOOR gegen Rheuma, Arthrose, Gelenk-, Frauen- und Prostataleiden  
 Hallenbad, Sauna, Dampfbad, Tennis - Thalasso-, Heu-, Aromabäder  
 Kurhaus Dr. Lauter - Kurhausstraße 81 - 82433 Bad Kohlgrub  
 Telefon (0 88 45) 97-0 Telefax (0 88 45) 97-2 59

[www.erlebe-gesundheit.de](http://www.erlebe-gesundheit.de)

**KLASSENFAHRTEN (incl. Transfer, Unterbringung, Programmgestaltung nach Absprache)**  
**NACH BERLIN**

Infobroschüre anfordern bei:  
 biss, Freiligrathstraße 3, 10967 Berlin, Tel. 0 30 / 6 93 65 30



**Deutsche Evangelische Oberschule Kairo**

Die Deutsche Evangelische Oberschule Kairo, eine vom Bundesverwaltungsamt in Köln -Zentralstelle für das Auslandsschulwesen- personell und finanziell unterstützte deutsche Auslandsschule,

sucht zum 1. Februar, spätestens zum 1. August 2004, Lehrkräfte mit Lehrbefähigung (1. und 2. Staatsexamen) in den Fächern

**Mathematik - Biologie - Chemie - Erdkunde - Englisch - Kunst - Musik**

Lehrkräfte, die wir direkt unter Vertrag nehmen, erhalten: Vertragsabschluss für zwei Jahre (Verlängerung möglich), Umzugsbeihilfe und jährliche Heimatflüge

Die Schule bietet Ihnen eine interessante und anspruchsvolle Aufgabe in einem arabischen Land. Zur Zeit werden etwa 1200 Schüler und Schülerinnen aus Ägypten und aus deutschsprachigen Ländern unterrichtet, welche unsere Schule vom Kindergarten bis zum Abitur besuchen können. Die Unterrichtssprache ist Deutsch.

Ihre vollständigen Bewerbungsunterlagen schicken Sie bitte an folgende Adresse:

**Deutsche Evangelische Oberschule Kairo  
 Auswärtiges Amt  
 Deutsche Botschaft Kairo / DEO  
 11020 Berlin**

Telefon: +20 2 748164 Fax: +20 2 7481648  
 Email: leiter@deokairo.de Homepage: www.deokairo.de