

D 85/2882+2

Max-Planck-Institut  
für Bildungsforschung  
Dokumentenabteilung  
1 Berlin 33, Lentzeallee 94  
① 85/2882+2

# Max-Planck-Institut für Bildungsforschung

Max Planck Institute for Human Development and Education

Jürgen Baumert, Peter Martin Roeder,  
Fritz Sang, Bernd Schmitz

Leistungsentwicklung und Ausgleich  
von Leistungsunterschieden in  
Gymnasialklassen

Nr. 12/SuU

Oktober 1985



Beiträge aus dem Forschungsbereich Schule und Unterricht  
Contributions from the Center for School Systems and Instruction



\*10075100\*

Jürgen Baumert, Peter Martin Roeder,  
Fritz Sang, Bernd Schmitz

Leistungsentwicklung und Ausgleich  
von Leistungsunterschieden in  
Gymnasialklassen

Nr. 12/SuU

Oktober 1985

Herausgegeben vom  
Forschungsbereich Schule und Unterricht  
Center for School Systems and Instruction

Max-Planck-Institut für Bildungsforschung  
Max Planck Institute for Human Development and Education  
Lentzeallee 94, D-1000 Berlin 33

Die „Beiträge“ aus den Forschungsbereichen sollen Arbeitspapiere und Forschungsergebnisse aus den einzelnen Arbeitsgruppen unabhängig von einer Veröffentlichung in Büchern oder Zeitschriften schnell zugänglich machen. Die Herausgabe erfolgt in der Verantwortung des jeweiligen Forschungsbereichs.

Papers in the „Contributions“ series are issued by the research centers at the Max Planck Institute for Human Development and Education to facilitate access to manuscripts regardless of their ulterior publication.

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, nur mit schriftlicher Genehmigung der Autoren.  
All rights reserved. No part of this paper may be reproduced without written permission of the authors.

Exemplare können angefordert werden bei  
Copies may be ordered from

**Max-Planck-Institut für Bildungsforschung**  
Lentzeallee 94, D-1000 Berlin 33

Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen

Achievement Growth and the Reduction of Achievement Differentials Within Classrooms

Zusammenfassung

Im gängigen Jahrgangsunterricht stehen Lehrer immer wieder vor dem Problem, Leistungsunterschiede zwischen den Schülern ausgleichen zu wollen, ohne zugleich leistungsstärkere Schüler zu benachteiligen. Wie schwierig diese Aufgabe ist, zeigen Untersuchungen zum Mastery Learning (beträchtlicher Bedarf an zusätzlicher Lernzeit) und zum Mathematikunterricht in der Hauptschule (Leistungsausgleich vorwiegend auf Kosten der leistungsstärkeren Schüler). Die vorliegende Untersuchung geht dem Problem des Verhältnisses von Leistungsförderung und Leistungsausgleich auf der Basis einer repräsentativen Stichprobe von ca. 400 Gymnasialklassen nach, deren Schulleistungen in den Fächern Deutsch, Englisch und Mathematik zu Beginn und am Ende des 7. Schuljahrs erhoben wurden und deren Lehrer nach ihren methodisch-didaktischen Orientierungen und Praktiken befragt wurden. Mit Hilfe geeigneter statistischer Verfahren wird der Versuch unternommen, einige der Kontextfaktoren und Unterrichtsstrategien zu bestimmen, die zu einer mehr oder weniger befriedigenden Lösung dieser Doppelaufgabe beitragen. Unter anderem wird geprüft, ob das von Dahllöf/Lundgren vorgeschlagene Modell der Orientierung des Lehrers an einer Steuerungsgruppe von Schülern zur Erklärung der empirischen Befunde ausreicht.

Summary

How can teachers equalize achievement within their classrooms without impeding the progress of high achieving students? Empirical studies on mastery learning seem to contradict the expectation that additional learning time for low achieving students can be reduced markedly while these students progress in their task. A survey on mathematics learning in German Secondary Modern Schools shows that a reduction in the dispersion of achievement scores within classrooms normally results in slower progress for high achieving students and no benefit for low achievers. The study presented here is based on a representative sample of seventh grade students in selective secondary schools (Gymnasien) who have been tested in three subjects at the beginning and at the close of the school year. Their teachers have answered a questionnaire concerning their teaching practices and objectives. The analysis focusses on the development over the school year of mean achievement scores, achievement scores of initially high and low achieving students, and achievement dispersion within classrooms. Differences between classrooms in these respects can be accounted for by a combination of contextual factors and teaching practices.



## Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen

### 1. Das allgemeine theoretische Problem

Nach dem Ausgleich von Leistungsunterschieden in Gymnasialklassen zu fragen, mag zunächst abwegig erscheinen. Denkt man an die wissenschaftliche und bildungspolitische Auseinandersetzung über Chancengleichheit und Chancenausgleich scheint die Fragestellung theoretisch und empirisch falsch plazierte zu sein. Im Bildungsprogramm der Gesamtschule etwa hat der Gedanke, Fähigkeits- und Leistungsunterschiede zu mindern, große Bedeutung. Ebenso gehört die Vorstellung kompensatorischer Förderung mittlerweile zum pädagogischen Selbstverständnis der Grundschule. Das Gymnasium dagegen gilt gerade als Vertreter eines Systems, das die Probleme zunehmender Leistungsunterschiede durch frühe Auslese und Verteilung der Schüler auf Bildungsgänge unterschiedlicher Niveaus und gerade nicht durch kompensatorische Bemühungen zu lösen sucht. Divergenzbegrenzung spielt danach im Gymnasium nur insofern eine Rolle, als fehlerhafte Übergangentscheidungen durch interne Auslese nachzubessern sind.

Jeder Kundige weiß natürlich, daß es sich bei dieser Darstellung um ein Zerrbild handelt. Schon der Hinweis auf den gleichmäßigen und erheblichen Rückgang der internen Auslese des Gymnasiums während der letzten Dekaden könnte allzu einseitige Vorstellungen korrigieren (vgl. Baumert 1980, S. 617, und 1985; Philipp/Witjes 1982). Auf diese Berichtigungen soll es uns jedoch hier nicht ankommen. Wir möchten uns zunächst einmal von pädagogischer Programmatik trennen, um den Zugang zu einem allgemeinen Strukturproblem des Unterrichts in Jahrgangsklassen zu eröffnen.

Der Unterricht in der Jahrgangsklasse beruht auf der Vorstellung, daß eine Altersgruppe, von der man annimmt, daß sie sich im wesentlichen auf einer Entwicklungsstufe befindet, gemeinsam von einem definierten obligatorischen Sockelniveau zum nächsten fortschreitet. Der Jahrgangsunterricht sieht im Grunde von inter- und intraindividuellen Unterschieden ab, so daß Gleichbehandlung von Schülern und optimale Förderung versöhnt werden. An diesem Punkt hat ja die Kritik an der Jahrgangsklasse immer wieder angesetzt (vgl. Furck 1961; Ingenkamp 1969). Überlagert wird dieses egalitäre Grundmuster der Jahrgangsklasse durch die Institutionalisierung des Leistungsgedankens: Schüler werden systematisch an Tüchtigkeitsmaßstäben gemessen, die zum Schulbeginn stärker ipsativer, später zunehmend überindividueller Art sind. Am sichtbarsten wird die Korrektur des Gleichheitsprinzips durch Billigkeitsvorstellungen an den Übergangsstellen zu weiterführenden Bildungsgängen, an denen knappe (oder knapp gehaltene) Chancen nach den Gesichtspunkten von Eignung und Leistung verteilt werden. Dennoch sind innerhalb der einzelnen Bildungsgänge - mögen sie den ganzen Jahrgang erfassen oder niveaumäßig differenziert sein - der Vergrößerung von Leistungsunterschieden Grenzen gesetzt, solange man von der Vorstellung des gemeinsamen Fortschritts der Altersgruppe und von Klassenunterricht nicht abgeht. Der Unterricht in der Jahrgangsklasse erzeugt gerade bei zunehmenden und durch die Leistungsbewertung verdeutlichten Fähigkeits- und Leistungsunterschieden einen strukturellen Bedarf an ausgleichenden Maßnahmen, die Bedürftigkeitsgesichtspunkte zur Geltung bringen. Die Möglichkeit der Klassenwiederholung widerspricht dem nicht. Die Einrichtung der Nichtversetzung berücksichtigt, daß in Einzelfällen Abweichungen auftreten können. Aber gerade als Ausnahme bestätigen sie die Regel - oder werden zum Ärgernis und zum Gegenstand schulaufsichtlichen Handelns. Jenseits aller Progammatik ist im Unterricht der Jahrgangsklasse das Problem eines Ausgleichs von "equality and excellence" strukturell angelegt.



## 2. Forschungslage

Im Zusammenhang der soziologischen Frage nach der Reproduktion gesellschaftlicher Ungleichheit lag ein besonderes Interesse der Erziehungswissenschaft auf schulischen Verteilungsmaßnahmen - seien es die Übergänge als den Gelenkstellen von Bildungskarrieren oder die Abschlüsse und die damit verbundenen Berechtigungen. Weniger Aufmerksamkeit dagegen wurde der Disparitätsentwicklung innerhalb der einzelnen Bildungsgänge zuteil, es sei denn, daß Versagerquoten zu einem Diskussionsthema wurden. Eine Ausnahme macht die Gesamtschule, deren Bildungsprogramm neben den Gedanken der Chancengleichheit ausdrücklich den des Chancenausgleichs setzt. Im Angebot optimaler Bildungs- und Entwicklungsmöglichkeiten für jeden einzelnen Schüler möchte die Gesamtschule Chancengleichheit verwirklichen, zugleich aber auch durch kompensatorische Maßnahmen außerschulisch bedingte Entwicklungsdivergenzen so weit ausgleichen, bis ein gemeinsames Sockelniveau schulischer Grundbildung erreicht ist. Damit erhält die Lösung des in jedem Schulklassenunterricht angelegten Problems der Verbindung von optimaler Qualifikation und Ausgleich von Leistungsunterschieden programmatische Bedeutung. Im Zusammenhang der Gesamtschulreform sind auch verschiedene empirische Arbeiten entstanden, die unter dem Gesichtspunkt der Unterrichtsdifferenzierung des Problems der Disparitätsminderung untersuchen (vgl. Morawietz 1980; Haußer 1981). Die Frage nach der Kompatibilität von optimaler individueller Förderung und Kompensation wird allerdings nur sehr selten ausdrücklich gestellt. Schaut man über die Landesgrenzen hinweg, so findet man - verständlicherweise - keine vergleichbar enge Verknüpfung des Problems der Disparitätsminderung mit schul- und unterrichtsorganisatorischen Fragen. In der angelsächsischen Forschung wird die Fragestellung vor allem im theoretischen Zusammenhang zielerreichenden Lernens und im Kontext von Programmen zur kompensatorischen Erziehung behandelt. Die in der Bundesrepublik entstandenen Arbeiten zur Auswirkung

unterrichtsorganisatorischer Maßnahmen lassen sich zwanglos einem weiteren Kreis von Untersuchungen zum "mastery learning" zuordnen (vgl. Block 1974; Bloom 1976).

Ansätze zielerreichenden Lernens beruhen im wesentlichen auf folgender gemeinsamer Grundstruktur: Zunächst wird der grundlegende Stoff einer in der Regel kurzen Unterrichtseinheit in der (heterogenen) Lerngruppe eingeführt. Nach einem Diagnosetest werden remediale Lerngruppen nach der Gesamtleistung oder nach speziellen Ausfällen gebildet, die in methodischer Abwandlung Teile des Fundamentums wiederholen, bis ein festgelegter Anteil von Schülern eine gemeinsame Basis des für die folgende Unterrichtseinheit notwendig erachteten Vorwissens erreicht hat. Dieses System ermöglicht, aus remedialen Erwägungen heraus zeitweilig homogenere Lerngruppen zu bilden oder individualisierten Unterricht zu erteilen, ohne für eine Gruppe leistungsschwächerer Schüler auf den Unterricht im anregungsreichen Milieu der heterogenen Klasse verzichten zu müssen. In seiner üblichen Form verbindet das Konzept zielerreichenden Lernens den Gedanken eines Unterrichts in wechselnden Lerngruppen, wie ihn schon Peter Petersen im Jena-Plan entwickelt hat, mit Vorstellungen "adaptiven Unterrichts" (vgl. Petersen 1927; Treiber/Petermann 1976; Cronbach/Snow 1977).

Dieses Modell beruht auf zwei theoretischen Grundannahmen über die Funktion der Zeit im Lernprozeß. Zunächst wird postuliert, daß der Lernfortschritt im wesentlichen eine Funktion der benötigten und verfügbaren Lernzeit sei (vgl. Carroll 1963; Wiley 1976; Harnischfeger/Wiley 1977; Treiber 1982). Dieses einflußreiche Carrollsche Lernmodell wird mit der Zusatzannahme verknüpft, daß infolge der systematischen Absicherung von Vorkenntnissen die Varianz der Lernrate zurückgehe, da sich der Bedarf an zusätzlicher Lernzeit auf seiten der langsameren Lerner schrittweise vermindere und die Lerngeschwindigkeiten sich aneinander angleichen (vgl. Garner 1978; Arlin 1984a).

Unterrichtsversuche mit diesem Modell zeigen, daß seine Realisierung curricular und organisatorisch ausgesprochen voraussetzungsvoll und auf Dauer schwer durchzuhalten ist (vgl. Eigler/Straka 1978). Dort, wo zielerreichendes Lernen jedoch praktiziert wird, erweist es sich traditionellem Schulunterricht im Hinblick auf ein erreichbares Sockelniveau überlegen. Die vorliegenden empirischen Befunde stimmen in diesem Punkt hinreichend überein (vgl. die Forschungsberichte von Block 1971; Block/Burns 1976; Burns 1979; Arlin 1984b). Auch die in der Bundesrepublik durchgeführten Untersuchungen zur flexiblen Differenzierung geben Hinweise, daß durch remediale Angebote die Leistungsstreuung im curricularen Basisbereich der Fächer Mathematik und Englisch verkleinert werden kann (vgl. Borchers/Krämer 1972; Vieth/Tönsing 1975; Vieth/Sommer 1976; Morawietz 1978). Ob man bei diesen Effekten freilich schon von zielerreichendem Lernen sprechen kann, ist fraglich; denn die Mittelwertdifferenzen unterschiedlicher Leistungsgruppen bleiben in der Regel auch nach der Differenzierungsphase im Bereich des Fundamentalstoffes beträchtlich (vgl. Morawietz 1978). Diese Befunde sind jedoch insofern besonders bedeutsam, da sie unter halbwegs realistischen Bedingungen hinsichtlich unterschiedlicher Zeitbudgets für langsamere und schnellere Lerner gewonnen wurden.

Während man davon ausgehen kann, daß durch die Gewährung zusätzlicher Lernzeiten für langsamere Lerner Leistungsdivergenzen vermindert werden können, ist die Annahme, daß der zusätzliche Zeitbedarf im Laufe zielerreichenden Unterrichts zurückgehe, umstritten. Die experimentellen Befunde zeigen zunächst, daß der zusätzliche Zeitbedarf langsamer Lerner, um Mastery-Bedingungen zu erfüllen, ganz erheblich ist. Je nach Gegenstand und Heterogenität der Kenntnisvoraussetzungen beträgt der Zusatzbedarf zwischen 40 und 70 Prozent der ursprünglich zur Verfügung gestellten Lernzeit. Relativiert man die Lernerträge am benötigten Zeitaufwand, so ist traditioneller Unterricht den Ansätzen zielerreichen-

den Lernens überlegen (vgl. Barr/Dreeben 1977; Arlin/Webster 1983; Arlin 1984b). Ferner legen Ergebnisse jüngerer Arbeiten, die gezielt die Frage der Lernratenveränderung untersuchen, den Schluß auf stabile Lernzeitunterschiede nahe (vgl. die Zusammenstellung bei Arlin 1984b). Bei Aufgabenstellungen, die langsame Lerner an die aktuellen Grenzen ihrer Leistungsfähigkeit führen, scheint überdies der Grenznutzen abzunehmen. Vermutlich läßt sich die Beziehung von Zeitaufwand und Lernertrag besser durch eine logarithmische als durch eine lineare Funktion approximieren (vgl. Fredrick/Walberg 1980; Wienold u.a. 1982; Aldridge 1983).

Auch wenn man von der Frage veränderlicher oder konstanter Lernraten absieht, ergibt sich allein aus dem anfänglich benötigten, erheblichen Zusatzbedarf an Lernzeit, der den in unseren Schulen üblichen Rahmen zusätzlicher Förderstunden weit überschreitet, das Problem von Wartezeiten für lernschnellere Schüler. Solange man von einem halbwegs vergleichbaren Zeitbudget für alle Schüler ausgeht, müssen die schnelleren Lerner während der remedialen Unterrichtsphasen sinnvoll beschäftigt werden, ohne daß zukünftiger gemeinsamer Unterricht vorweggenommen wird. Das Auftreten von Wartezeiten ist freilich kein besonderes Problem differenzierten Unterrichts, der unter der Zielsetzung von Chancenausgleich erteilt wird. Wartezeiten gibt es ebenso im Unterricht in der Jahrgangsklasse, die auch hier durch die Vorstellung vom gleichen Lernfortschritt hervorgerufen werden (vgl. Jackson 1968; Smith/Geoffrey 1968; Doyle 1977). Bei Differenzierungsmaßnahmen wird dieses Strukturproblem jedoch unübersehbar und verlangt nach geplanten Lösungen. Die bisherigen empirischen Lösungen sind insofern nicht befriedigend, als zwar ein höheres gemeinsames Sockelniveau erreicht wird, aber schnellere Lerner die Wartezeit nicht optimal nutzen und so in ihren Fortschritten gebremst werden (vgl. Resnick 1977; Arlin 1984b). Wenn die experimentelle Anordnung in einem für alle Schüler gleichen Zeitrahmen die Selbstregulierung des Lerntempo erlaubt, tritt auch bei zielerreichendem Unterricht

keine Divergenzminderung ein, sondern die Leistungsstreuung nimmt teilweise sehr schnell zu (vgl. Atkinson 1968; Arlin/Westbury 1976; Hyman/Cohen 1979). Diese Befunde legen die Vermutung nahe, daß Varianzminderung bei chancenausgleichendem Unterricht zumindest teilweise durch geringere Fortschritte leistungsstärkerer Schüler zustandekommt. In ähnliche Richtung weisen auch Ergebnisse der Untersuchungen, die Fend an Gesamtschulen durchführte. Eine Verringerung außerschulisch bedingter Disparitäten war nur bei Minderung des Leistungsniveaus der sozial begünstigten Schülergruppe zu verzeichnen (vgl. Fend 1982; Wottawa 1982; Baumert u.a. 1985).

Ein weiteres Unterrichtskonzept, von dem man annimmt, daß es besonders zur kompensatorischen Förderung sozial benachteiligter Schüler geeignet sei, ist die sogenannte "direkte Instruktion". Was im angelsächsischen Sprachraum mit diesem Begriff bezeichnet wird, hat sehr viel Ähnlichkeit mit dem in der Bundesrepublik zumindest in den Lehrgangsfächern noch immer vorherrschenden Muster lehrergeleiteten Unterrichts. Diese Unterrichtsform ist durch hohe Instruktionsintensität, relativ klare Sachstruktur, vom Lehrer ausgehende Interaktion mit der ganzen Klasse, Unterbrechungsarmut, enge Kontrolle sowie gezieltes Üben und Wiederholen gekennzeichnet (vgl. Rosenshine 1976 und 1979; Good 1979; Fend 1984). Die Mehrzahl der empirischen Arbeiten zu diesem Unterrichtskonzept wurde an amerikanischen Grundschulen durchgeführt und beschäftigt sich insbesondere mit dem Erwerb grundlegender Kulturtechniken, teilweise mit besonderer Konzentration auf Schulen in sozial benachteiligten Gebieten. Diese Untersuchungen belegen im wesentlichen die Überlegenheit direkter Instruktionsformen gegenüber einem Unterricht, in dem Wechsel der Sozialform (Gruppenarbeit, Partnerarbeit und Stillarbeit) üblich ist, Lehrer einzelne Schüler individuell betreuen und Schüler freier miteinander kommunizieren oder gar ihre Aufgaben selbst suchen. Ihre Befunde konnten im großen und ganzen in experimentellen Studien, die auch

Sekundarschulen miteinbezogen, bestätigt werden (vgl. die Übersichten bei Brophy 1979 und 1982; Good 1979; Borich/Fenton 1977; Treiber 1981). Soweit die Studien Interaktionseffekte zwischen Schülervoraussetzungen und Unterrichtsgestaltung untersuchten, ergaben sich Hinweise, daß leistungsschwächere und ängstliche Schüler im besonderen Maße von einem stark strukturierenden Unterricht profitieren (vgl. Good/Power 1976; Ebmeier/Good 1979; Anderson/Scott 1978).

Obwohl die mittlerweile umfangreiche Forschung zur direkten Instruktion von Programmen zur kompensatorischen Erziehung ausging, vollzog unseres Wissens keine der Untersuchungen den Schritt zu einer simultanen Analyse von Divergenz- und Leistungsentwicklung in Lerngruppen. Selbst systematische Prüfungen auf Wechselwirkungen zwischen Schülermerkmalen und Instruktionsform sind relativ selten. Zwar gab es Anregungen, bei der Untersuchung von Schul- und Unterrichtseffekten über den Vergleich von Mittelwerten hinauszugehen und andere Verteilungsparameter in der Analyse zu berücksichtigen. Klitgaard (1975) etwa diskutiert solche Möglichkeiten in seinem Aufsatz "Going Beyond the Mean...". Brown und Saks (1975 und 1978) gehen einen Schritt in diese Richtung, wenn sie im theoretischen Rahmen eines Produktionsmodells den Einfluß der Mittelverteilung im Schulwesen auf Streuung und Niveau von Schulleistungen untersuchen. Entschiedener wenden sich dem Problem der mehrkriterialen Analyse Baker und McPartland (1984) zu, die in einem Extremgruppenvergleich unterschiedlich effektiver Schulen dem Zusammenhang von Niveau und Variation von Schulleistung nachgehen (vgl. auch Klitgaard/Hall 1977). Arbeiten, die das Problem von Divergenzminderung und optimaler Qualifikation unter den Bedingungen üblichen Klassenunterrichts behandeln, sind jedoch rar. Zu diesen Arbeiten gehört die Untersuchung, die Weinert und andere zum Mathematikunterricht in Hauptschulklassen der 5. Jahrgangsstufe durchführten (vgl. Treiber u.a. 1976). Ihre Befunde geben Anlaß, die bisherigen Forschungsergebnisse zur kompensatorischen Wirkung direkten

Unterrichts innerhalb von heterogenen Lerngruppen zu differenzieren. Die Daten dieser Studie besagen, daß zwar auch lernschwächere Schüler von Formen der direkten Instruktion profitieren, aber wider Erwarten gerade die leistungsstärkeren den instruktionsintensiven Unterricht besser nutzen, so daß bei dieser Unterrichtsführung Leistungsunterschiede zunehmen. Die gemeinsame Analyse von Streuungs- und Leistungsentwicklung zeigt die Gegenläufigkeit beider Effekte. Ein Ausgleich von Disparitäten wurde im wesentlichen durch Minderung des Leistungszuwachses schnellerer Lerner erzielt, ohne daß langsamere Schüler nennenswerte Gewinne verzeichneten (vgl. Treiber 1980; Treiber/Schneider 1980; Treiber u.a. 1982; Treiber/Weinert 1985).

Die Resultate beider Forschungsprogramme machen die Probleme divergenzmindernder Unterrichtsformen ausreichend deutlich. Die Arbeiten, die zu Ansätzen zielerreichenden Lernens entstanden, verweisen bei flexibler Unterrichtsorganisation auf die von langsamen Lernern benötigten zusätzlichen Unterrichtszeiten und bei Organisationsformen, die Differenzierungsmaßnahmen nur innerhalb eines für alle Schüler gemeinsamen Zeitrahmens erlauben, auf die für schnellere Schüler entstehenden Wartezeiten, in denen "vertiefender", aber nicht fortschreitender Unterricht erteilt werden muß. Überträgt man dies auf die Situation der heterogenen Jahrgangsklasse, so verlangt eine chancenausgleichende Unterrichtsführung eine "akzeptable" Balance zwischen diagnostischen und remedialen Maßnahmen einerseits und einem befriedigenden Anspruchsniveau und ausreichender Instruktionsintensität andererseits. Die Ergebnisse, die Treiber und andere über die Effekte direkter Instruktion im Mathematikunterricht berichten, runden dieses Bild ab. Nach ihren Befunden kann man von einem instruktionsintensiven Unterricht - jedenfalls unter den Bedingungen einer Hauptschulklasse - fördernde Wirkung, aber keine Varianzreduktion und von einer stärkeren Akzentuierung remedialer Elemente Divergenzminderung, aber kaum Förderung erwarten.

Da die verfügbaren empirischen Daten sich entweder auf Innovationsprogramme oder auf Regelunterricht in Einrichtungen beziehen, denen man besondere Bemühungen um Chancenausgleich zumindest zuschreibt, ist die Generalisierbarkeit der Befunde einigermaßen ungewiß. Wenn freilich die Balancierung von Leistungs- und Divergenzentwicklung ein allgemeines Strukturproblem des Unterrichts in der Jahrgangsklasse ist, sollten wir bei Untersuchungen unter anderen Rahmenbedingungen ein vergleichbares Ergebnismuster erwarten. Eine Überprüfung dieser These ließe sich mit der Untersuchung divergenzmindernden Unterrichts an einer Schulform vornehmen, in deren Bildungsprogramm der Gedanke des Chancenausgleichs hinter dem der Förderung der Leistungsbesten deutlich zurücktritt. Im folgenden wollen wir diese Prüfung anhand der Sekundäranalyse einer Studie über Determinanten der Leistungsentwicklung in Gymnasialklassen vornehmen.

### 3. Spezifizierung der Fragestellungen

Geht man vom traditionellen Bildungsanliegen des Gymnasiums aus, die leistungsfähigsten Schüler eines jeden Jahrgangs in einem einheitlichen Bildungsgang zur Hochschulreife zu führen, so haben weder der Gedanke der internen Differenzierung noch der des Chancenausgleichs größere Bedeutung. Dem scheint die Erziehungsphilosophie der Gymnasiallehrer zu entsprechen, wenn im Widerstreit von Fördern und Auslesen der Auslese häufiger als in anderen Schulformen der Vorrang gegeben wird (vgl. Philipp/Witjes 1982). Diese Entsprechung von Programm und Einstellung legt die Vermutung nahe, daß ein varianzreduzierender Unterricht im Gymnasium ausgesprochen selten anzutreffen sei und das Problem zunehmender Heterogenität der Jahrgangsklasse zufriedenstellend über Klassenwiederholungen gelöst werden könne, ohne daß eine streuungsreduktive Unterrichtsführung erforderlich werde.



So einleuchtend diese Annahme im ersten Augenblick erscheint, so sprechen doch verschiedene theoretische Überlegungen und empirische Hinweise gegen sie. Zunächst ist die Klassenwiederholung keineswegs eine leicht zu handhabende Maßnahme zur Begrenzung der Heterogenität einer Klasse. Sie greift überhaupt nur bei Leistungsmängeln in mehreren Fächern, ist von der Übereinstimmung des Kollegiums abhängig und grundsätzlich als justiziable Ausnahme definiert. Die Klassenwiederholung bietet wohl die Möglichkeit, die Rahmenbedingungen des Unterrichts langfristig in gewissen Grenzen zu beeinflussen, verändert jedoch an den Bedingungen des täglichen Unterrichts in einer mehr oder minder heterogenen Lerngruppe nichts. Die kollektive Situation des Klassenunterrichts verlangt vom Lehrer ganz bestimmte soziale Ordnungsleistungen, deren Vernachlässigung Unterricht beenden kann. Schon um die Mindestvoraussetzungen von Unterricht zu sichern, nämlich eine Situation zu schaffen, in der Störungen und Unterbrechungen so weit eingegrenzt sind, daß zielgerichtetes Arbeiten überhaupt möglich wird, erfordert eine Unterrichtsführung, die eine Balance zwischen Überforderung und Langeweile anstrebt. Doyle sowie Barr und Dreeben weisen wohl zu Recht darauf hin, daß Lehrer grundlegende Unterrichtsentscheidungen, die Anspruchsniveau und Instruktionstempo betreffen, zunächst gar nicht unter Gesichtspunkten fachlichen Lernens treffen, sondern von dem naheliegenden Ziel geleitet werden, eine möglichst gleichmäßige "Kooperation" ihrer Schüler zu erreichen (vgl. Doyle 1979; Barr/Dreeben 1977; Cusick 1983).

Erfahrene Lehrer verstehen es, bestimmte, für Außenstehende unauffällige Verhaltensweisen von Schülern als Warnhinweise zu deuten und mit präventiven Reaktionen einen störungsarmen Unterrichtsablauf und ein relativ gleichmäßiges Aufmerksamkeitsniveau zu erreichen. So beantworten Lehrer etwa Anzeichen von nachlassender Aufmerksamkeit gleichsam automatisch durch besondere Konzentration auf diese Schüler oder Schülergruppen (vgl. Doyle 1979; Arlin 1979). Solche "cues" seitens

der Schüler sind wahrscheinlich die wichtigsten Hinweise zur Anpassung des Schwierigkeitsgrades von Aufgabenstellungen und zur Regulierung des Instruktionstempos. Wann und wieviel neuer Stoff einzuführen ist, ob eine schwierigere Variante behandelt werden kann oder welches Maß an Veranschaulichung, Übung und Wiederholung erforderlich ist, darüber entscheiden Lehrer im einzelnen in der konkreten Unterrichtssituation. Es spricht einiges dafür, daß diese balancierenden Unterrichtsentscheidungen überwiegend reaktiv auf Schülersignale erfolgen und nur selten proaktive Maßnahmen darstellen (vgl. Gump 1967). Interviews, die Arlin und Barr mit Lehrern über ihre Unterrichtsgestaltung führten, ist zu entnehmen, daß Gruppierungsentscheidungen sehr bewußt und gezielt getroffen werden, während über Instruktionstempo und Themenwechsel kaum Auskünfte zu erhalten sind (vgl. Arlin 1979; Barr 1975). Bei zunehmender Heterogenität der Lerngruppe scheint es insbesondere für unerfahrene Lehrer schwieriger zu werden, in habitualisierter Weise angemessen zu reagieren (vgl. Sanford 1980; Evertson u.a. 1981).

Dahllöf und Lundgren haben zu zeigen versucht, daß Lehrer Instruktionstempo und Fortschreiten im Stoff an einer "Steuerungsgruppe" orientieren, zu der sie die Schüler zählen, deren Fähigkeitsmaße etwa zwischen dem 10. und dem 25. Perzentil liegen (vgl. Dahllöf 1971; Lundgren 1972). Auf das Problem, daß Steuerungsgruppen empirisch weder durch Lehrerbefragung noch durch Unterrichtsbeobachtung ohne weiteres zu identifizieren sind, das Modell aber gleichwohl von einer hohen Rationalität der Lehrerentscheidungen ausgeht, ist gelegentlich hingewiesen worden (vgl. Westbury 1973; Arlin 1979). Gleichwohl stellt die "framing-Theorie" Dahllöfs einen furchtbaren Ansatz dar, da sie die situativen Unterrichtsbedingungen ausdrücklich berücksichtigt. Ebenso scheint in der Verortung einer Steuerungsgruppe im unteren Drittel der Leistungsverteilung der erwägenswerte Gedanke angelegt zu sein, daß Unterrichtsentscheidungen von Lehrern sowohl durch das Leistungsniveau als auch durch die Leistungsstreu-

ung einer Lerngruppe beeinflußt werden. Das Modell vereinfacht und rationalisiert diesen Prozeß vermutlich jedoch zu sehr.

Wir gehen davon aus, daß Lehrer ihre Entscheidungen in einem komplexeren, aber zugleich diffuseren Bezugsrahmen treffen, der in hohem Maße eingeschliffene Verhaltensabläufe zuläßt. Das Bezugssystem wird unserer Annahme nach durch zwei Parameter bestimmt: Einmal orientieren sich Lehrer in ihren Unterrichtsentscheidungen an einer globalen Einschätzung des Leistungsniveaus einer Lerngruppe. Es gibt Hinweise, daß diese Einschätzung Lehrern zuverlässiger gelingt, als man gemeinhin annimmt und daß sich Lehrer über solche Einschätzungen auch verständigen können (vgl. Roeder u.a. 1985). Sodann berücksichtigen sie jedoch auch die Fähigkeitsstreuung einer Lerngruppe, über die sie durch eine Vielzahl von Schülerreaktionen im Unterricht Auskunft erhalten. Wenn eine zunehmende Leistungsvarianz die Aufrechterhaltung eines einigermaßen gleichmäßigen Aufmerksamkeitsniveaus zu gefährden scheint, reagieren Lehrer störungspräventiv mit streuungsmindernden Maßnahmen. Situationsdeutung und Reaktion sind wahrscheinlich besonders beim erfahrenen Lehrer weitgehend habitualisiert, so daß der Ereignisfluß des Unterrichts kaum unterbrochen wird. Wir vermuten, daß vor allem auf Übungs- und Wiederholungsformen zurückgegriffen wird, die als Lernhilfen gelten und zugleich alle Schüler der Klasse gleichmäßig beschäftigen. In einem zweidimensionalen Bezugsrahmen hat man also mit einem Netzwerk von sich wechselseitig beeinflussenden situativen Bedingungen und Lehrerreaktionen zu rechnen (vgl. Good/Power 1976; Treiber 1981).

Im Rahmen dieser theoretischen Überlegungen gehen wir von folgenden Untersuchungsannahmen aus:

- Auch an ausleseorientierten Schulformen wie dem Gymnasium ist Unterricht anzutreffen, der Leistungsunterschiede mindert.

- Divergenzminderung und Leistungsentwicklung verhalten sich gegenläufig.
- Bei streuungsverringender Unterrichtsführung verlangsamten sich die Lernfortschritte von Schülern im oberen Leistungsbereich.
- In ihrer Unterrichtsgestaltung berücksichtigen Lehrer die jeweilige Klassensituation; dabei reagieren sie auf Höhe und Streuung der Vorkenntnisse.
- Als streuungsmindernde Maßnahmen kommen vor allem Abstufungen der Aufgabenschwierigkeit, Verlangsamung des Unterrichtstempos und Verringerung der angebotenen Stoffmenge sowie die Intensivierung von Übung und Wiederholung in Frage - also Reaktionsformen, die zum üblichen Verhaltensrepertoire von Lehrern zählen.

#### 4. Untersuchungsanlage und Variablenbeschreibung

Empirische Grundlage der Untersuchung ist ein Datensatz, der im Rahmen des "Projekts Schulleistung" im Schuljahr 1969/70 an einer für die Länder der Bundesrepublik und Berlin West repräsentativen Stichprobe von Gymnasialklassen der 7. Jahrgangsstufe erhoben wurde (vgl. Edelstein 1970)<sup>1</sup>. Die Untersuchung bezog sich auf die Schulfächer Deutsch, Englisch und Mathematik. Die Leistungs- und Persönlichkeitsdaten der Schüler wurden zu zwei Zeitpunkten, am Anfang und am Ende des 7. Schuljahres, erhoben. Zum zweiten Meßzeitpunkt wurden die Fachlehrer der untersuchten Klassen anhand eines umfangreichen, jeweils fachspezifischen Fragebogens um Auskünfte über ihre Unterrichtsstrategien gebeten (vgl. Hopf 1980). Es liegen verarbeitbare Datensätze für 12.594 Schüler aus 427 Klassen und für insgesamt 1.130 Deutsch-, Englisch- und Mathematiklehrer vor.

Für unsere Analysen wurden folgende Variablengruppen ausgewählt:

Soziales Milieu: Zur Beschreibung der häuslichen Lernumwelt wurden ein Index, der aus den höchsten Bildungsabschlüssen von Mutter und Vater gebildet wurde, verwendet, ferner die Häufigkeit verschiedener bildungsrelevanter Tätigkeiten (Lesen, Malen, Musizieren, naturkundliche Beschäftigungen), die tägliche Fernsehzeit sowie die häusliche Belastung durch schulfremde Aufgaben.

Intelligenz: Das allgemeine intellektuelle Fähigkeitsniveau der Schüler wurde mit dem IST-Amthauer bestimmt.

Fachleistungen: Die Schülerleistungen wurden in den Fächern Deutsch, Englisch und Mathematik mit zwei Tests zu Beginn und zum Ende der 7. Jahrgangsstufe erhoben (Test A und Test B). Die Testaufgaben wurden auf der Grundlage von Stofferhebungen entworfen und entsprachen den Lehrplananforderungen und dem tatsächlichen Unterrichtspensum für die 7. Klasse des Gymnasiums (vgl. Edelstein u.a. 1968; Zeiher 1972). Die internen Konsistenzen der verwendeten Gesamttests liegen zwischen  $r = .82$  und  $.91$ .

Unterrichtsvariablen: Im Anschluß an jüngere Entwicklungen der Lehr-Lern-Forschung wurden auf der Grundlage der Befragung von Fachlehrern zu ihren Unterrichtsstrategien eine Reihe von Skalen und Kennziffern zur Quantität und Qualität der Instruktion gebildet. Die in unserer Analyse verwendeten Indikatoren sind mit ihren Kenngrößen im Anhang beschrieben.

Die folgenden Analysen werden auf Klassenebene durchgeführt. Zur Kennzeichnung, wie sich die Dispersion der Schülerleistungen in den einzelnen Schulklassen zwischen den Meßzeitpunkten entwickelte, wurde für jede Klasse der Quotient aus den Standardabweichungen der Messung B und A berechnet. Diesem Indikator gaben wir gegenüber anderen Kennziffern vornehmlich aus theoretischen Gründen den Vorzug<sup>2</sup>.

Zur Beschreibung der Leistungsentwicklung der 7. Gymnasialklassen verwenden wir die auf Klassenebene gemittelten Residuen der multiplen Regression von Testleistung B auf Testleistung A, die gemessene Intelligenz und Merkmale der sozialen Herkunft sowie der häuslichen Lernumwelt. Die Residuen dienen als ein um die Effekte individueller Voraussetzungen "bereinigtes" Maß für den Erfolg der gymnasialen Unterrichtsbemühungen während eines Schuljahres<sup>3</sup>. Ob mit vorgelagerten Einflüssen der Länderzugehörigkeit zu rechnen ist, wurde anhand einer Teilstichprobe mit einem hierarchisch geschachtelten varianzanalytischen Design geprüft. Während sich bei Querschnittanalysen bedeutsame Ländereffekte zeigen, werden sie bei den residualisierten Testwerten entweder nicht mehr signifikant oder sind von geringer praktischer Bedeutung. Für unsere Untersuchung, die nur Residuen verwendet, verzichten wir deshalb auf die Berücksichtigung dieser Analyseebene.

## 5. Ergebnisse

### 5.1 Divergenzmindernder Unterricht im Gymnasium

Es gibt wenig Zweifel, daß Ende der sechziger Jahre die Mehrzahl der Gymnasiallehrer - bei aller Differenziertheit des Urteils - ihre Schule als Einrichtung verstand, die vor allem der Förderung und Auslese der Begabten verpflichtet sei. Der Gedanke, Unterschiede auszugleichen und langsame Lerner auf ein befriedigendes Leistungsniveau zu führen, spielt in ihrer Erziehungsphilosophie kaum eine Rolle, wie Schefer und Zeiher in ihren Untersuchungen zum Gesellschaftsbild und pädagogischen Selbstverständnis des Gymnasiallehrers belegen (vgl. Schefer 1969; Zeiher 1973). Die Ergebnisse Zeihers beruhen auf Daten der "Schulleistungsstudie", die uns zur Reanalyse zur Verfügung stehen. Mehr als drei Viertel der Befragten befürworteten eine strenge Auslese und gut die Hälfte argumentierte so mit Entschiedenheit. Daß man diesen Befund nicht einfach auf die Gegenwart

übertragen darf, versteht sich von selbst. Zwar scheinen Gymnasiallehrer auch zum Ende der siebziger Jahre weniger Probleme mit dem Dilemma von Fördern und Auslesen zu haben als ihre Kollegen an anderen Schulformen (vgl. Bauer 1980; Engelhardt 1982). Hinsichtlich des Urteils über die Bildbarkeit von Schülern sind die Befunde jedoch nicht konsistent (vgl. Bauer u.a. 1979; Kischkel 1979). Vor gut 15 Jahren aber war in den Gymnasien noch ein Erziehungsklima anzutreffen, das pädagogische Vorstellungen von einem chancenausgleichenden Unterricht gewiß nicht begünstigte.

Freilich sind nicht nur deshalb Unterrichtsformen, die es erlauben, flexibler auf Schülerunterschiede einzugehen, nach unserer Datenlage selten. Die pädagogische Diskussion über wechselnde Sozialformen, über informelle Tests und Diagnostik, über die Struktur von Fachleistungen und mögliche Förderung hatte ja damals gerade erst eingesetzt. So ist es nicht verwunderlich, wenn binnendifferenzierende oder individualisierende Maßnahmen, ein vielfältiger Medieneinsatz oder Bemühungen, Schülerinitiativen zu wecken, Ausnahmererscheinungen sind. (Zu den Häufigkeitsverteilungen einzelner Vorgehensweisen im Mathematikunterricht vgl. Hopf 1980). Aber auch bis heute scheinen sich die Verhältnisse im Mathematikunterricht nicht grundlegend gewandelt zu haben, wenn man die Ergebnisse, die eine Replikation von Teilen dieses Fragebogens zu Unterrichtsstrategien mit Lehrern der gymnasialen Oberstufe erbracht hat, einmal auf die Mittelstufe überträgt oder an die vergleichbaren Befunde aus Hauptschulen denkt (vgl. Reiss 1983; Treiber/Weinert 1985).

Geht man von dieser Situationsbeschreibung aus, ist es überraschend, daß ein streuungsmindernder Unterricht in den 7. Klassen des Gymnasiums sogar in den traditionellen Hauptfächern Deutsch, Englisch und Mathematik keineswegs unüblich ist. Tabelle 1 ist zu entnehmen, daß je nach Unterrichtsfach in zwischen einem und zwei Drittel der untersuchten Klassen

die Leistungsstreuung zwischen beiden Testzeitpunkten abnimmt. Im Englischunterricht tritt dieser Effekt in 149 von 415 (35,9 Prozent) und im Deutschunterricht sogar in 278 von 415 (67,1 Prozent) Klassen auf. Der Mathematikunterricht nimmt eine Mittelstellung ein; hier ist in 44,4 Prozent der Klassen ein streuungsreduktiver Unterricht anzutreffen. Ob streuungsintensiv oder streuungsreduktiv unterrichtet wird, ist unabhängig vom Alter der Lehrer und ihrer Einstellung zu Fördern und Auslesen.

## 5.2 Leistungs- und Streuungsentwicklung

Vergleicht man die residualisierten Testleistungen von Klassen mit zu- und abnehmender Leistungsdivergenz zunächst für das Fach Mathematik, so läßt sich der Befund der Heidelberger Hauptschuluntersuchung bestätigen, daß sich Disparitätsminderung und Leistungszuwachs gegenläufig verhalten (vgl. Treiber 1980; Treiber/Weinert 1985). Die Mittelwerte beider Gruppen unterscheiden sich signifikant (vgl. Tabelle 1 und 2). Allerdings ist der Effekt der Gruppenzugehörigkeit mit einem  $\omega^2 = .05$  gering. Die Korrelation zwischen den Indikatoren für Leistungs- und Streuungsveränderung liegt mit  $r = .24$  erheblich unter dem von Treiber berichteten Zusammenhang in Hauptschulklassen ( $r = .50$ )<sup>4</sup>. Ein Blick auf die Ergebnisse von Schülergruppen mit unterschiedlichen Leistungsvoraussetzungen (gemessen am Test A) zeigt, daß die Klassenmittelwerte gruppenspezifische Effekte verdecken. Bei divergenzmindernder Klassenführung sinken die Testwerte im oberen Leistungsdrittel beträchtlich, ohne daß Schüler im unteren Leistungsdrittel Gewinne zu verzeichnen hätten. Die Mittelwertunterschiede in der Leistungsspitze nähern sich einer Standardabweichung und sind damit nicht mehr trivial. Die Divergenzentwicklung erklärt nunmehr auch 17 Prozent der Varianz der Testwerte der leistungsstärksten Gruppen der untersuchten Klassen.



In den sprachlichen Fächern findet sich das gleiche Grundmuster, allerdings in abgeschwächter Form. Im Saldo unterscheiden sich Klassen mit zunehmender und abnehmender Streuung nicht signifikant; die vorhandenen Mittelwertunterschiede weisen allerdings in die gleiche Richtung wie im Fach Mathematik. Ein Disparitätsausgleich kommt im muttersprachlichen und fremdsprachlichen Unterricht durch leichte Verbesserung von Schülern im unteren und deutliche Verschlechterung der Klassenkameraden im oberen Leistungsdrittel zustande. Die Mittelwertdifferenzen liegen für die Leistungsspitze bei etwa einer Dreiviertel-Standardabweichung, während sie für das untere Leistungsdrittel unter einer halben Standardabweichung bleiben. Einen Überblick über die korrelativen Zusammenhänge gibt Tabelle 3.

### 5.3 Gruppierung nach Leistungs- und Divergenzentwicklung

Ein Leistungsunterschiede ausgleichender Unterricht ist also im Gymnasium auch in den ausleserelevanten und streng benoteten Fächern Mathematik, Deutsch und Englisch keine Ausnahme (vgl. Roeder u.a. 1965; Roeder u.a. 1985). Bemühungen, die Leistungsschere in der Jahrgangsklasse sich nicht allzu weit öffnen zu lassen, sind offenbar kein Spezifikum von Schulformen mit besonderem Förderungsprogramm. Disparitätsminderung und optimale Förderung aller Schüler sind jedoch auch im Gymnasium nicht ohne weiteres zu vereinbaren. Die Verbindung beider Ansprüche stellt ein Optimierungsproblem dar, das von Klasse zu Klasse und von Lehrer zu Lehrer unterschiedlich gelöst wird. Um einen Überblick über die empirische Verteilung unterschiedlicher Lösungsmuster zu gewinnen, wurden die untersuchten Klassen nach Divergenz- und Leistungsentwicklung (zunehmend/abnehmend beziehungsweise überdurchschnittlich/unterdurchschnittlich) gruppiert. Die Verteilung der Klassen auf die vier Gruppen sowie die mittleren Leistungswerte für die jeweiligen Klassen insgesamt und für

zwei Teilgruppen mit unterschiedlichen Leistungsvoraussetzungen weisen die Tabellen 4 und 5 getrennt nach Unterrichtsfächern aus. Deutschlehrern gelingt es am häufigsten, varianzreduktiv zu unterrichten und zugleich überdurchschnittlich zu qualifizieren. Etwa 30 Prozent der Klassen weisen diese Merkmalskombination auf (Gruppe 3), während 18 Prozent der Deutschlehrer hohe Leistungszuwächse bei zunehmenden Leistungsunterschieden erzielen (Gruppe 1). In den Lehrgangsfächern Englisch und Mathematik ist der Anteil der Klassen im Quadranten 3 nur etwa halb so groß; in diesen Fächern scheinen hohe Qualifikationsleistungen häufiger mit Streuungszunahme einherzugehen (etwa ein Drittel der Klassen befindet sich jeweils in Gruppe 1). Bei den unterdurchschnittlich qualifizierenden Klassen (Gruppe 2 und 4) ergeben sich im Hinblick auf die Häufigkeit varianzvergrößernder beziehungsweise -mindernder Unterrichtsführung Schwankungen von Fach zu Fach.

Im Hinblick auf die durchschnittliche Leistungsentwicklung, die in qualifikationsintensiven Klassen erreicht wird, unterscheidet sich ein streuungsmindernder Unterricht in keinem Fach von streuungsvergrößernden Vorgehensweisen (Gruppe 1 und 3), wie die t-Werte der Kontraste zeigen. Dasselbe gilt auch für unterdurchschnittlich qualifizierende Klassen (Gruppe 2 und 4) mit Ausnahme des Faches Mathematik. Hier ist eine signifikante Interaktion zwischen Streuungs- und Leistungsentwicklung zu verzeichnen. Im Mathematikunterricht erzielen divergenzmindernd und zugleich mit unterdurchschnittlichem Lernerfolg unterrichtete Klassen die ungünstigsten Leistungsresultate.

Hinter den vergleichbaren Leistungsergebnissen qualifikationsintensiver Klassen mit unterschiedlicher Streuungsentwicklung (Gruppe 1 und 3) verbergen sich wiederum differenzierte Muster für Gruppen mit unterschiedlichen Leistungsvoraussetzungen. In Klassen, in denen die Leistungsunter-

schiede zunehmen, erreicht das obere Leistungsdrittel deutlich bessere Lernerfolge, während in den streuungsreduktiv geführten Klassen Schüler am entgegengesetzten Ende der Leistungsverteilung profitieren (das gleiche Bild wiederholt sich im wesentlichen in Klassen mit ungünstiger Leistungsentwicklung (Gruppe 2 und 4). Also auch in Schulklassen mit über Erwartungen gutem Lernerfolg stehen sich bei unterschiedlicher Streuungsentwicklung jeweils Erträge und Kosten gegenüber. Beide Unterrichtsformen stellen freilich gegenüber einem Unterricht, der die jeweiligen Vorzüge vereint, noch keine optimale Lösung dar. In der Tat lassen sich in jedem Fach Klassen finden - es sind immer etwa 10 Prozent der Stichprobe -, in denen Lehrern dies gelingt. Zu dieser Gruppe gehören jeweils ihrem Anteil in der Stichprobe entsprechend sowohl Klassen mit zunehmender als auch mit abnehmender Streuung.

Wenn unsere Annahme zutrifft, daß die Balancierung von Leistungs- und Divergenzentwicklung ein generelles Strukturproblem des Schulklassenunterrichts ist, das sich je nach mittlerem Kenntnissniveau und Leistungsstreuung einer Klasse in unterschiedlicher Schärfe stellt, müssen die skizzierten empirischen Lösungen des Optimierungsproblems von der jeweils spezifischen Kenntnissituation in den unterrichteten Klassen beeinflusst sein. So sollten etwa zwischen Klassen mit unterschiedlicher Divergenzentwicklung Unterschiede in der Vorkenntnisstreuung und zwischen streuungsreduktiv unterrichteten Klassen mit unterschiedlicher Leistungsentwicklung ebenso Unterschiede im Vorkenntnissniveau nachweisbar sein.

Ein Blick auf die Tabellen 6 und 7, die eine Beschreibung der Kenntnissituation in den unterschiedlich unterrichteten Klassen und eine Zusammenfassung der Ergebnisprüfungen auf Mittelwertunterschiede bieten, zeigt, daß diese Vermutungen im wesentlichen bestätigt werden können. In allen Unterrichtsfächern ist die Streuung im Vortest in Klassen mit

abnehmender Leistungsstreuung deutlich höher (signifikanter Haupteffekt des Faktors Divergenzentwicklung). Im Fach Englisch und Mathematik unterscheiden sich die Gruppen 3 und 4 (divergenzmindernder Unterricht mit unterdurchschnittlichem Qualifikationserfolg) erwartungsgemäß auch im mittleren Niveau der verfügbaren Vorkenntnisse (signifikanter Haupteffekt des Faktors Leistungsentwicklung beziehungsweise signifikante Interaktion). Nur im Fach Deutsch wird der vorhandene Niveauunterschied, der zwar in die erwartete Richtung weist, noch nicht als Haupt- oder Interaktionseffekt statistisch bedeutsam. Ein ausreichend hohes Vorkenntnisniveau scheint in der Tat eine günstige Voraussetzung zu sein, auch bei hoher Kenntnisstreuung über Erwarteten gute Lernerfolge zu erzielen. Am deutlichsten wird dies im Fall der "Optimalklassen", deren mittleres Vorkenntnisniveau im Fach Mathematik etwa eine halbe Standardabweichung über dem Gesamtmittel liegt. Fallen niedriges Vorkenntnisniveau und zugleich hohe Leistungsstreuung zusammen, entsteht eine ungewöhnlich schwierige Lernsituation, der Lehrer vermutlich nicht selten mit einem übungintensiven und wenig stimulierenden Unterricht begegnen<sup>5</sup>.

#### 5.4 Situative Bedingungen und Unterrichtsstrategien am Beispiel des Faches Mathematik

Im folgenden soll am Beispiel des Mathematikunterrichts zu zeigen versucht werden, mit welchen Unterrichtsstrategien Lehrer auf Folgeprobleme heterogener Leistungsvoraussetzungen antworten beziehungsweise auf Anzeichen möglicher Folgeprobleme präventiv reagieren. Besonders ethnographische Studien zu Schule und Unterricht haben das Tempo der sozialen Interaktion, die Unvorhersehbarkeit der Ereignisfolgen sowie die Gleichzeitigkeit oft unverträglicher Geschehnisse im Unterricht als strukturelle Merkmale der kollektiven Situation beschrieben und betont, in welchem

Maße Lehrer Reaktionen habitualisieren müssen, schon um hinreichend entlastet zu sein, damit sie in entscheidenden "Gelenkstellen" des Unterrichts und Schullebens bewußt und überlegt entscheiden können (vgl. Smith/Geoffrey 1968; Metz 1978; Cusick 1973; Doyle 1977 und 1979). Wir vermuten, daß gerade die Reaktionen auf immer wiederkehrende Unterrichtssituationen - und das sind nicht zuletzt jene Situationen, in denen Lehrer auf die Leistungsheterogenität ihrer Klassen gestoßen werden - besonders bei erfahrenen Lehrern in einer Form eingeschliffen sind, in der sie sich möglichst ohne Bruch in den Gang des Unterrichts einfügen (vgl. Kounin 1970). Zu Maßnahmen, die divergenzmindernd wirken könnten und zugleich solche Anforderungen erfüllen, dürften vor allem Verlangsamung des Unterrichtstempos, Verringerung der dargebotenen Stoffmenge, Abstufung der Aufgabenschwierigkeit, Darbietung anschaulicher Beispiele sowie die Intensivierung von Übung und Wiederholung zählen.

Um zu einer Beschreibung des Unterrichts in Klassen mit unterschiedlicher Streuungsentwicklung zu gelangen (wobei zunächst die Leistungsentwicklung unberücksichtigt bleibt), wurde eine schrittweise Diskriminanzanalyse mit verschiedenen Indizes und Skalen, die Merkmale "direkter Instruktion" sowie remediale Unterrichtspraktiken erfassen, als diskriminierenden Variablen gerechnet. Mit einer Auswahl von sieben Variablen wurde eine vertretbare Lösung erreicht. Zu diesen Variablen gehören das mittlere Vorkenntnisniveau und die Streuung im Vortest als Rahmenbedingungen sowie jeweils ein Indikator zur Quantität der Instruktion, zum Anspruchsniveau, zur Übungs- und Wiederholungsintensität, Präsentation des Unterrichtsstoffes und zur argumentativen Begründung von Noten und Zensuren. Die Diskriminanzfunktion wurde anhand einer Teilstichprobe von 206 Klassen geschätzt (vgl. Tabelle 8). Bei einer Reklassifikation werden 68 Prozent der Fälle korrekt gruppiert. Dieses Ergebnis bleibt auch bei einer Kreuzvalidierung (70 Prozent richtig gruppierter Fälle) stabil.

Unterricht in Klassen mit zunehmender und abnehmender Streuung läßt sich anhand der Gruppenzentroide etwa folgendermaßen beschreiben: In Klassen, deren Leistungsstreuung im Laufe des Untersuchungsjahres zunimmt, findet Unterricht unter eher günstigen Ausgangsbedingungen statt. Die Vorkenntnisse der Schüler streuen weniger und liegen auf einem relativ hohen Niveau. Die Lehrer dieser Klassen können das Unterrichtstempo entsprechend hoch ansetzen und den Schülern mehr Stoff anbieten. Die Lehrer dieser Klassen betonen weder in besonderem Maße ihr Anspruchsniveau noch bemühen sie sich ungewöhnlich intensiv um klare Präsentation an schwierigen Stellen oder um Übung und Wiederholung. Gelegentlich begründen sie ihre Noten und Urteile argumentativ. Divergenzmindernder Unterricht dagegen ist unter eher schwierigen Voraussetzungen anzutreffen. In diesen Klassen streuen die Vorkenntnisse bei tendenziell niedrigerem Niveau stark. Das Instruktionstempo wird dementsprechend zugunsten von Übungsphasen zurückgenommen. Dennoch versuchen die hier unterrichtenden Lehrer, ein angemessenes Anspruchsniveau aufrechtzuerhalten und der ungünstigen Situation unter anderem durch Bemühung um besonders klare Präsentation und eindeutige Urteilsmaßstäbe gerecht zu werden.

Um diese Gegenüberstellung zu differenzieren, wurden auch die nach Divergenz- und Leistungsentwicklung gruppierten Klassen (vgl. Tabellen 4 und 5) anhand von Merkmalen der Vorkenntnissituation und der Unterrichtsgestaltung zu trennen versucht. Die bei der ersten Analyse verwendete Auswahl von Diskriminationsvariablen wurde um drei weitere Variablen erweitert: In der auf Klassenebene gemittelten Note für Mitarbeit kommt eine allgemeine Einschätzung von Lehrern zum Ausdruck, wie angenehm in dieser Klasse Unterricht zu erteilen sei. Zwei weitere Skalen erfassen, inwieweit Unterricht durch das Lehrerbuch strukturiert wird

und in welchem Ausmaß Basiskenntnisse behandelt wurden. Für die Gruppentrennung wurden zwei Diskriminanzfunktionen verwendet, deren Gewichte anhand einer Substichprobe von 190 Fällen geschätzt wurden. Ihre Kennwerte sowie die standardisierten Faktorkoeffizienten der abhängigen Variablen sind der Tabelle 9 zu entnehmen. Bei einer mit Hilfe beider Funktionen vorgenommenen Reklassifikation werden 51 Prozent der Fälle richtig gruppiert. Bei einer Kreuzvalidierung sinkt dieser Anteil auf 47 Prozent. Die Fehlklassifikationen sind vor allem Folge unbefriedigender Trennung der Gruppen 2 und 3 von Gruppe 1. Knapp 30 Prozent der Klassen aus Gruppe 2 und 40 Prozent der Klassen aus Gruppe 3 werden fälschlicherweise in Gruppe 1 plazierte. Wahrscheinlich ließe sich die Diskrimination der Gruppen verbessern, wenn uns Mediationsvariablen auf Lehrerseite, etwa die Wahrnehmung von Leistungsniveau und Leistungsstreuung, zur Verfügung ständen.

Abbildung 1 veranschaulicht anhand der Zentroide die Lokalisierung der Klassen im zweidimensionalen Diskriminanzraum. Die erste Dimension wird durch den bipolaren Faktor "Üben von Basisfertigkeiten" versus "hohes Anspruchsniveau und hohes Unterrichtstempo" beschrieben, während die zweite Dimension durch den Kontrast von hoher Streuung und eher niedrigem Leistungsniveau einerseits und hohem Niveau und geringer Streuung andererseits gekennzeichnet ist. Überdurchschnittlich qualifizierend, aber zugleich Unterschiede vergrößernd wirkt ein Mathematikunterricht, in dem Lehrer die guten Vorkenntnisse und geringe Kenntnisstreuung der Klasse nutzen und ein hohes Anspruchsniveau einhalten, den Stoff zügig durchnehmen und an besonders schwierigen Stellen klar und strukturiert präsentieren (Gruppe 1). In Klassen der Gruppe 2 verstehen Lehrer ähnlich günstige Voraussetzungen nicht aufzugreifen. Hier werden Schüler, deren Vorkenntnisse relativ hoch und homogen sind, mit einem anspruchslosen und wenig strukturierten Unterricht

konfrontiert, von dem alle wenig Gewinn haben, mit dem leistungsstärkere Schüler aber offenbar noch besser zu-recht kommen. Vermutlich findet man in diesen Klassen in besonderem Maße mißlingenden Unterricht, weil Bedingungen elementarer Klassenführung nicht erfüllt sind (vgl. Good 1978; Good u.a. 1978). Von diesen Lehrern sind jene Lehrer, die in den Klassen der Gruppe 4 Unterricht erteilen, abzusetzen. Denn diese Lehrer reagieren mit ihrem eher elementaren und repetitiven Unterricht auf eine schwierige Ausgangssituation, die durch ein geringes Vorkenn-nisniveau und zugleich große Kenntnisunterschiede gekenn-zeichnet ist. Der Unterricht in den Klassen der Gruppe 3 kann offenbar - darauf deutet schon die hohe Zahl der Fehlklassifikationen hin - sehr unterschiedlich sein. Wenn eine hohe Leistungsstreuung bei gutem durchschnitt-lichen Kenntnisniveau vorliegt, scheint sich die Unter-richtsgestaltung sehr der in den Klassen der Gruppe 1 an-zunähern. Es wird anspruchsvoll und zügig unterrichtet. Bei abnehmendem Kenntnisniveau dagegen gleicht sich der Unterricht mit stärkeren remedialen Akzenten eher dem der Gruppe 4 an.

### 5.5 Modellvergleich

Im folgenden soll versucht werden, den Zusammenhang von situativem Bezugsrahmen, Unterrichtsgestaltung und den beiden Instruktionseffekten - Leistungs- und Disparitäts-entwicklung - in einem Strukturgleichungsmodell für den gymnasialen Mathematikunterricht abzubilden und zu testen. In diesem Rahmen wollen wir die Dahllöfsche Annahme, daß sich Lehrer in ihren Unterrichtsentscheidungen an einer Steuerungsgruppe orientieren (Modell A), mit unserer Ver-mutung eines zweiparametrischen Bezugsrahmens konfrontie-ren (Modell B). Zum Zwecke des Vergleichs wählen wir ein vereinfachtes Gesamtmodell mit drei latenten Unterrichts-



variablen, die als Übungsintensität, Quantität der Instruktion und Anspruchsniveau zu bezeichnen sind. Als Situationsmerkmale gehen in das Gesamtmodell Leistungsstreuung und Leistungsniveau einer Schulklasse und zur Kennzeichnung der Steuerungsgruppe das Niveau des unteren Leistungsdrittels einer Klasse ein. Als Kriterien dienen Divergenz- und Leistungsentwicklung. Dieses Gesamtmodell schließt also die Modelle A und B ein. Die Abbildung 2 zeigt das Strukturgleichungsmodell mit seinen postulierten Zusammenhängen. Über die in der Abbildung wiedergegebenen Einflüsse hinaus wurden Meßfehlerkorrelationen zugelassen - und zwar für die Meßfehler der latenten Situationsvariablen, da sie auf ein und derselben Messung beruhen, sowie für die Meßfehler der Unterrichtsvariablen, die die verwendeten Stoffe erfassen.

Die Modellprüfung wurde mit dem Programm Lisrel V (vgl. Jöreskog/Sörbom 1981) vorgenommen. Das Programm erlaubt den direkten Vergleich der konkurrierenden Annahmen (A und B) im Rahmen eines Gesamtmodells.  $\chi^2$ -Test ( $\chi^2 = 36.7$ ;  $df = 27$ ;  $p = 0.10$ ), Goodness-of-Fit-Index (.95) und Residuen (RMSR = .033) sprechen für eine befriedigende Anpassung des Gesamtmodells. Ein Blick auf die Strukturgleichungskoeffizienten zeigt zunächst den auffälligen Befund, daß keiner der Koeffizienten, die Einflüsse der Steuerungsgruppe repräsentieren, signifikant wird. Werden diesem Befund entsprechend diese Strukturgleichungsparameter auf Null fixiert, so ergibt sich nach dem Vergleich der  $\chi^2$ -Tests keine Verschlechterung der Modellanpassung ( $\chi^2$  für die Differenz der beiden Modelle = 9.10;  $df = 5$ ). Setzt man dagegen zum Zwecke des expliziten Vergleichs zwischen Dahllöfs Modell und dem hier vorgeschlagenen die von Leistungsstreuung und Leistungsniveau ausgehenden Einflüsse auf Null - als situativer Einfluß bleibt dann allein die Steuerungsgruppe bestehen - so ist dieses Modell mit den empirischen Daten nicht mehr kompatibel ( $\chi^2 = 189.77$ ;

df = 34; p = 0.00). Die  $\chi^2$ -Differenz zwischen dem vollständigen und diesem reduzierten Modell ist hoch signifikant ( $\chi^2 = 153.0$ ; df = 7). Der direkte Vergleich der Steuerungsgruppentheorie mit unserer Annahme eines komplexeren Referenzrahmens spricht also für eine Revision des Dahllöfschen Konzepts. Ein Zweikomponentenmodell ist mit den uns zur Verfügung stehenden Daten deutlich besser verträglich. In Abbildung 2 werden der Übersichtlichkeit wegen nur die Koeffizienten des reduzierten Modells (B) wiedergegeben.

In diesem Modell liegt die multiple Determination für die Divergenzentwicklung bei .44 und für die Leistungsentwicklung bei .59. Trotz der Höhe dieser Koeffizienten ist das Modell jedoch keineswegs befriedigend. Die Verknüpfung der Leistungsstreuung mit Unterrichtsvariablen und Disparitätentwicklung gelingt nur begrenzt. Der theoretisch erwartete Einfluß zunehmender Heterogenität einer Klasse auf die Quantität der Instruktion ist nicht nachweisbar, und der Zusammenhang zwischen Streuung und Übungsintensität ist relativ schwach. Das hohe Gewicht des direkten Einflusses der Leistungsstreuung auf die Divergenzentwicklung ist unter anderem auch Ausdruck der im Hinblick auf die Unterrichtsvariablen unzureichenden Spezifikation unseres Modells. An diesem Punkte werden auch Grenzen der Erhebungsmethode deutlich.

Faßt man die Informationen, die das Strukturmodell vermittelt, zusammen, ergibt sich folgendes Bild: Lehrer stimmen ihren Unterricht hinsichtlich des allgemeinen Anspruchsniveaus, Unterrichtstempos und Stoffangebots sowie der Übungsintensität auf situative Bedingungen der Schulklasse ab. In besonderem Maße orientieren sie sich offenbar an einer globalen Einschätzung des Leistungsniveaus ihrer Klasse; auf ungünstige Leistungsvoraussetzungen reagieren sie vor allem mit vermehrter Übung und Verringerung der

Stoffmenge. Einer zunehmenden Leistungsstreuung versuchen sie mit erhöhter Übungsintensität zu begegnen. Häufiges Üben und langsames Fortschreiten im Stoff führen zwar zu einer Disparitätsminderung, hemmen aber zugleich die Leistungsentwicklung. Rechnet man unsere Analysen getrennt für Schüler des oberen und unteren Leistungsdrittels, so wiederholt sich das berichtete Grundmuster. Im einzelnen zeigen sich jedoch charakteristische Besonderheiten: So erweisen sich die Quantität der Instruktion und das allgemeine Anspruchsniveau für die Entwicklung der leistungsstärkeren Schüler als deutlich wichtiger. Ein intensives Üben von Basisfertigkeiten während des Unterrichts ist jedoch entgegen dem remedialen Anspruch für beide Gruppen gleichermaßen wenig hilfreich.

#### 6. Zusammenfassung und Diskussion

Ein Leistungsunterschiede ausgleichender Unterricht ist auch im selektiven Gymnasium der ausgehenden sechziger Jahre eine übliche Unterrichtsform. Das gilt selbst für die auslesebedeutsamen Fächer Deutsch, Englisch und Mathematik. Im Fach Englisch etwa ist dieser Unterricht in gut einem Drittel und im Fach Deutsch sogar in zwei Dritteln der Klassen anzutreffen. Divergenzminderung und Leistungsentwicklung verhalten sich auch im Gymnasium tendenziell gegenläufig; ein Ausgleich von Leistungsunterschieden ist nicht ohne weiteres mit optimaler Qualifikation zu vereinbaren. In nach Leistungsgruppen getrennten Analysen konnte gezeigt werden, daß bei streuungsverringendem Unterricht erhebliche Leistungseinbußen im oberen Leistungsdrittel relativ schmalen Gewinnen im unteren Leistungsdrittel gegenüberstehen. Gleichwohl ließen sich in jedem Fach etwa 10 Prozent der Klassen identifizieren, deren Lehrern es gelingt, die Vorzüge beider Unterrichtsformen zu verbinden.

Im Vergleich zu den Befunden der Heidelberger Hauptschuluntersuchung ist der Zusammenhang von Divergenz- und Leistungsentwicklung jedoch gering. Im Gymnasium gelingt es offenbar besser als in der Hauptschule, den Ausgleich von Leistungsunterschieden mit Leistungsförderung zu verbinden. Dieses zunächst überraschende Ergebnis verliert an Auffälligkeit, wenn man sich Situation und Struktur der Hauptschule vergegenwärtigt. Die Hauptschule ist eine Einrichtung, die in den vergangenen 20 Jahren unter zweiseitigen Druck geraten ist, da sie einmal als Reformschule weiterführende Bildungswege erschließen sollte und zugleich von der Abwanderung ihrer Stammklientel betroffen wurde. Auf die Populationsverschiebung reagierte ein Teil der Hauptschullehrer mit verstärkten sozialpädagogischen Bemühungen, die - wie Fallstudien belegen - die unterrichtlichen Aufgaben der Schule in den Hintergrund drängen können (vgl. Roeder 1984). Hier hat man wohl in erster Linie mit einer Verringerung der Quantität der Instruktion zu rechnen. Zugleich darf man jedoch nicht übersehen, daß die Hauptschule trotz eines entgegengesetzten Vorurteils - immer noch eine sehr heterogene Schülerschaft betreut und mit internen Differenzierungsmaßnahmen versucht, ihr gerecht zu werden. Die interne Selektivität der Hauptschule ist in der Vergabe unterschiedlicher Abschlüsse und Berechtigungen, in manchen Bundesländern auch der Fachleistungsdifferenzierung, der Klassenwiederholung oder dem Angebot berufsbefähigender Lehrgänge für eine schwierige Minderheit der Schüler institutionell abgesichert. Das System dieser Schule scheint von der Unterrichtsorganisation her sogar eher auf Förderung von Leistungsdivergenzen als auf Disparitätsminderung angelegt zu sein. Gerade Lehrer mit einem hohen Bildungsanspruch, die zumindest einen Teil ihrer Schüler zu einem qualifizierten Abschluß führen wollen, sind auch im Anfangsunterricht der Hauptschule in stärkerem Maße als im Gymnasium gezwungen, eine zunehmende Heterogenität der Schulklasse zu tolerieren

und mit ihr erfolgreich umzugehen. Der unterschiedliche Zusammenhang von Leistungs- und Dispersionsentwicklung in Hauptschule und Gymnasium ist vermutlich in erster Linie Ausdruck von System- und Populationsunterschieden.

Ob im Gymnasium divergenzmindernd oder divergenzvergrößernd unterrichtet wird, ist unabhängig vom Alter der Lehrer und ihrer Einstellung zur Begabtenauslese. Unsere Befunde stützen die Vermutung, daß hier Reaktionen auf unterschiedliche Arbeitssituationen in den Schulklassen zum Ausdruck kommen. Anhand eines vereinfachten Unterrichtsmodells für den gymnasialen Mathematikunterricht ließ sich zeigen, daß Lehrer ihren Unterricht auf situative Bedingungen der Schulklasse abstimmen, wobei sowohl Leistungsniveau als auch Leistungsstreuung der Schulklasse Orientierungspunkte darstellen. Mit dem Konzept einer Steuerungsgruppe, wie Dahllöf vorgeschlagen hat, werden diese Referenzgesichtspunkte nur unbefriedigend erfaßt. Das Anspruchsniveau des unterrichtenden Lehrers, die Quantität der Instruktion sowie die Intensität von Übung und Wiederholung erwiesen sich für die Leistungs- und Divergenzentwicklung als wichtige Prädiktoren. Ein hohes Anspruchsniveau und zügiges Fortschreiten im Stoff wirken zwar divergenzsteigernd, beeinflussen aber - jedenfalls im Mathematikunterricht - in überraschender Weise die Leistungsentwicklung von Schülern des oberen und unteren Leistungsdrittels positiv. Wider Erwarten wirkt eine ausgesprochen remediale Unterrichtsführung, die dem Üben und Wiederholen auch von Grundfertigkeiten besonders breiten Raum gibt, für beide Gruppen ungünstig.

## 7. Anhang

### 7.1 Anmerkungen

- 1 Die "Schulleistungsstudie" wurde als interdisziplinäres Projekt am Max-Planck-Institut für Bildungsforschung von folgenden Personen durchgeführt: M. von Cranach, W. Edelstein, H.-L. Freese, D. Hopf, E. Meischner, F. Sang, W. Stegelmann, H. J. Zeiher, H. Zeiher. Für ihr Einverständnis, die Daten sekundäranalytisch nutzen zu dürfen, danken wir ihnen.
- 2 Die verschiedentlich vorgeschlagene Steigung der interklassischen Regression von Test B auf Test A als "Entkopplungsindikator" ist inhaltlich schwer zu interpretieren, und jede Trennung von vermeintlich divergenzsteigernder oder divergenzmindernder Klassenführung bleibt beliebig (vgl. Treiber 1980; Burstein 1981; Treiber/Schneider 1980; Achtenhagen 1981). Zudem bietet dieser Indikator keine meßtechnischen Vorteile; die Folgen unvermeidlich kleiner Stichproben und die Ausreißerempfindlichkeit treffen auch dieses Konzept. Die gelegentlich empfohlene Korrelation zwischen allgemeiner Fähigkeit und Testleistung ist bei näherer Betrachtung der vorliegenden empirischen Ergebnisse wahrscheinlich eher eine Approximation der Instruktionsintensität als ein Indikator für Dispersionsveränderung (vgl. Simons u.a. 1975; Weinert u.a. 1980; Sörensen 1985). Ökonomische Disparitätsmaße schließlich, wie Gini-Koeffizient oder Lorenz-Kurve, bieten bei annähernd normalverteilten Variablen keine Vorteile; sie sind eher angemessen zur Kennzeichnung extrem schiefer Verteilungen (vgl. Allison 1978; Schwartz/Winship 1979; Baker/McPartland 1984).
- 3 Die gemittelten Residuen stellen für unseren Datensatz kovarianzanalytisch adjustierten Mittelwerten gleichwertige Meßgrößen dar. Die Korrelation zwischen beiden Maßen, die für eine Substichprobe von 100 Schulklassen berechnet wurde, liegt bei  $r = .99$ . Aus rechentechnischen Gründen werden die gemittelten Residuen bevorzugt (vgl. Keesling/Wiley 1974; Wiley 1976; Cronbach/Webb 1975; McPherson 1981). Das bei allen hierarchisch geschachtelten Daten, so auch bei Veränderungsmessungen an intakten Schulklassen auftretende Problem der Abhängigkeit der Beobachtungen (bei Analysen auf der individuellen Ebene) beziehungsweise der Heterogenität der klasseninternen Regressionssteigung (bei Klassenanalysen) wird bei Untersuchungen, die Leistungs- und Dispersionsentwicklung gleichzeitig behandeln, unübersehbar; gleichwohl bietet sich eine probate Lösung des Dilemmas bislang nicht an (vgl. Burstein 1981). Eine Clusterung von Schulklassen mit homogenen Regressionssteigungen verbietet sich, da damit gleichzeitig die

Varianz der Unterrichtsvariablen und der Divergenzentwicklung erheblich reduziert wird, so daß die Wahrscheinlichkeit, erwartbar schwache Effekte zu entdecken, unzuträglich gemindert wird.

- 4 Verwendet man, um die Koeffizienten besser vergleichbar zu machen, wie Treiber die Steigung der Intraklassenregression von Test B auf Test A als "Entkopplungsindikator", sinkt die Korrelation auf  $r = .21$ .
- 5 Wird bei Veränderungsmessungen auf Klassenebene aggregiert, ist bei ausreichender Testreliabilität das Problem einer Regression zur Mitte, soweit diese meßfehlerbedingt ist, nicht schwerwiegend. Allerdings kann man einen Regressionseffekt aufgrund der "Kumulation unwahrscheinlicher Ereignisse" (vgl. Furby 1973), die nur einzelne Klassen betreffen, nicht ausschließen. Bei Zweipunktmessungen gibt es dafür keine Korrekturmöglichkeiten. Allerdings sprechen unsere Befunde, wie sie in den Tabellen 4 und 5 wiedergegeben werden, nicht für Regressionseffekte (vgl. Baltès u.a. 1972; Furby 1973; Nesselroade u.a. 1980). Auch bei Veränderungswerten von Streuungsmaßen - hier der Quotient der Standardabweichungen - kann man die Wirksamkeit eines Regressionseffektes nicht ausschließen, wenn ein Zwischenklassenmeßfehler in Wechselwirkung mit dem Vortestniveau steht. Um den möglichen Fehlereinfluß einer Regression zur Mitte abzuschätzen, wurden die Analysen mit einer Auswahl von Klassen wiederholt, deren Vorteststreuung im mittleren Bereich liegt. Die Ergebnisse zeigten keinerlei Abweichungen.

Tabelle 1: Testleistungen (gemittelte Residuen) in Klassen mit zunehmender und abnehmender Leistungsstreuung nach Fächern und Schülergruppen mit unterschiedlichem Vorkenntnisniveau

Divergenz- entwicklung	Mathematik (n = 383)	Deutsch (n = 415)	Englisch (n = 415)
zunehmend	n = 213 (55,6 %)	n = 137 (32,9 %)	n = 266 (64,1 %)
insgesamt	M = .83 SD = 3.95	M = .38 SD = 2.54	M = .19 SD = 3.08
oberes Leistungsdrittel	M = 8.04 SD = 4.46	M = 7.12 SD = 2.88	M = 5.32 SD = 3.42
unteres Leistungsdrittel	M = -5.81 SD = 4.15	M = -5.63 SD = 2.89	M = -5.03 SD = 3.34
abnehmend	n = 170 (44,4 %)	n = 278 (67,1 %)	n = 199 (35,9 %)
insgesamt	M = -1.01 SD = 3.99	M = -.16 SD = 2.86	M = -.32 SD = 3.16
oberes Leistungsdrittel	M = 3.93 SD = 4.62	M = 4.49 SD = 3.33	M = 2.64 SD = 3.55
unteres Leistungsdrittel	M = -5.75 SD = 3.87	M = -4.31 SD = 3.03	M = -3.52 SD = 3.34



Tabelle 2 : Zusammenfassung der Varianzanalysen auf Klassenebene mit residualisierter Testleistung als Kriterium und Divergenzentwicklung (zunehmend/abnehmend) als Faktor

Fach	Analyseeinheit	df	F	p	$\omega^2$
Deutsch	Klasse insgesamt	1/413	3.63	.06	-
	oberes Leistungsdrittel	1/413	62.64	.00	.13
	unteres Leistungsdrittel	1/413	18.20	.00	.04
Englisch	Klasse insgesamt	1/413	2.60	.11	-
	oberes Leistungsdrittel	1/413	56.39	.00	.12
	unteres Leistungsdrittel	1/413	19.27	.00	.04
Mathematik	Klasse insgesamt	1/381	20.76	.00	.05
	oberes Leistungsdrittel	1/381	77.70	.00	.17
	unteres Leistungsdrittel	1/381	0.02	.89	-

Tabelle 3: Produkt-Moment-Korrelationen zwischen Divergenzentwicklung und residualisierter Testleistung nach Unterrichtsfächern und Vorkenntnissen

Fach	Testleistung		
	Klasse insgesamt	Oberes Leistungsdrittel	Unteres Leistungsdrittel
Deutsch	.11	.47	-.29
Englisch	.08 n.s.	.41	-.26
Mathematik	.24	.48	-.06 n.s.

Tabelle 4 : Gruppierung der Schulklassen nach Leistungs- und Divergenzentwicklung in den Fächern Deutsch und Englisch

Deutsch (n = 415)

Divergenzentwicklung	Leistungsentwicklung	
	überdurchschnittlich	unterdurchschnittlich
zunehmend	Gruppe 1	Gruppe 2
	n = 75 (18 %)	n = 62 (14,9 %)
	M (insgesamt) = 2.26	M (insgesamt) = -1.89
	M (oberes Drittel) = 9.07	M (oberes Drittel) = 4.77
	M (unteres Drittel) = -3.73	M (unteres Drittel) = -7.94
abnehmend	Gruppe 3	Gruppe 4
	n = 126 (30,4 %)	n = 152 (36,6 %)
	M (insgesamt) = 2.33	M (insgesamt) = -2.23
	M (oberes Drittel) = 6.95	M (oberes Drittel) = 2.45
	M (unteres Drittel) = -2.10	M (unteres Drittel) = -6.13

Englisch (n = 415)

Divergenzentwicklung	Leistungsentwicklung	
	überdurchschnittlich	unterdurchschnittlich
zunehmend	Gruppe 1	Gruppe 2
	n = 137 (33 %)	n = 129 (31,1 %)
	M (insgesamt) = 2.55	M (insgesamt) = -2.31
	M (oberes Drittel) = 7.62	M (oberes Drittel) = 2.87
	M (unteres Drittel) = -2.93	M (unteres Drittel) = -7.26
abnehmend	Gruppe 3	Gruppe 4
	n = 65 (15,7 %)	n = 84 (20,2 %)
	M (insgesamt) = 2.48	M (insgesamt) = -2.48
	M (oberes Drittel) = 5.50	M (oberes Drittel) = 0.45
	M (unteres Drittel) = -1.02	M (unteres Drittel) = -5.47

Gruppenvergleiche:

Klasse insgesamt	$t_{(1/3)}$	= -.27
	$t_{(2/4)}$	= 1.38
oberes Leistungsdrittel	$t_{(1/3)}$	= 6.32*
	$t_{(2/4)}$	= 6.70*
unteres Leistungsdrittel	$t_{(1/3)}$	= -5.12*
	$t_{(2/4)}$	= -5.49*

Klasse insgesamt	$t_{(1/3)}$	= -.24
	$t_{(2/4)}$	= .66
oberes Leistungsdrittel	$t_{(1/3)}$	= 5.65*
	$t_{(2/4)}$	= 7.00*
unteres Leistungsdrittel	$t_{(1/3)}$	= -5.00*
	$t_{(2/4)}$	= -5.10*

Tabelle 5: Gruppierung der Schulklassen nach Leistungs- und Divergenzentwicklung im Fach Mathematik

Gesamtstichprobe (n = 383)

Divergenzentwicklung	Leistungsentwicklung	
	überdurchschnittlich	unterdurchschnittlich
zunehmend	Gruppe 1	Gruppe 2
	n = 122 (31,8 %)	n = 91 (23,7 %)
	M (insgesamt) = 3.45	M (insgesamt) = -2.69
	M (oberes Drittel) = 10.62	M (oberes Drittel) = 4.58
	M (unteres Drittel) = -3.44	M (unteres Drittel) = -8.99
abnehmend	Gruppe 3	Gruppe 4
	n = 65 (16,9 %)	n = 105 (27,4 %)
	M (insgesamt) = 3.00	M (insgesamt) = -3.49
	M (oberes Drittel) = 8.16	M (oberes Drittel) = 1.32
	M (unteres Drittel) = -2.24	M (unteres Drittel) = -7.93

Ohne Extremgruppen (n = 250)<sup>1</sup>

Divergenzentwicklung	Leistungsentwicklung	
	überdurchschnittlich	unterdurchschnittlich
zunehmend	Gruppe 1	Gruppe 2
	n = 86 (34,4 %)	n = 53 (21,2 %)
	M (insgesamt) = 3.52	M (insgesamt) = -2.84
	M (oberes Drittel) = 10.76	M (oberes Drittel) = 4.77
	M (unteres Drittel) = -3.31	M (unteres Drittel) = -9.29
abnehmend	Gruppe 3	Gruppe 4
	n = 44 (17,6 %)	n = 67 (26,8 %)
	M (insgesamt) = 2.74	M (insgesamt) = -3.72
	M (oberes Drittel) = 7.73	M (oberes Drittel) = 1.00
	M (unteres Drittel) = -2.49	M (unteres Drittel) = -8.27

Gruppenvergleiche:

$$\begin{cases} t_{(1/3)} = 1.24 \\ t_{(2/4)} = 2.33^* \end{cases}$$

$$\begin{cases} t_{(1/3)} = 4.92^* \\ t_{(2/4)} = 6.97^* \end{cases}$$

$$\begin{cases} t_{(1/3)} = -2.67^* \\ t_{(2/4)} = -2.52^* \end{cases}$$

Vergleich beider Verteilungen (Gesamtstichprobe und Teilstichprobe ohne Extremgruppen)  $\chi^2_{(3)} = 1.32$  n.s.

$$\begin{cases} t_{(1/3)} = 1.77 \\ t_{(2/4)} = 2.03^* \end{cases}$$

$$\begin{cases} t_{(1/3)} = 5.03^* \\ t_{(2/4)} = 6.31^* \end{cases}$$

$$\begin{cases} t_{(1/3)} = -1.53 \\ t_{(2/4)} = -1.92 \end{cases}$$

<sup>1</sup> Nur Klassen mit einer im mittleren Bereich liegenden Standardabweichung im Test A.

Tabelle 6 : Vortestleistung und Standardabweichung in Klassen mit unterschiedlicher Leistungs- und Divergenzentwicklung

Divergenzentwicklung	Leistungsentwicklung					
	überdurchschnittlich		unterdurchschnittlich			
zunehmend	Gruppe 1		Gruppe 2			
	Deutsch:	M = 56.07 SD = 11.15	Deutsch:	M = 56.35 SD = 10.48		
	Englisch:	M = 54.09 SD = 9.43	Englisch:	M = 49.42 SD = 9.27		
	Mathematik:	M = 46.50 SD = 11.25	Mathematik:	M = 46.55 SD = 10.88		
abnehmend	Gruppe 3		Gruppe 4			
	Deutsch:	M = 54.00 SD = 12.15	Deutsch:	M = 52.85 SD = 11.68		
	Englisch:	M = 52.19 SD = 10.55	Englisch:	M = 48.80 SD = 10.07		
	Mathematik:	M = 48.39 SD = 12.78	Mathematik:	M = 44.18 SD = 12.21		

Tabelle 7: Zusammenfassung der Varianzanalysen auf Klassenebene mit dem Kriterium Vortestleistung bzw. Standardabweichung im Vortest und den Faktoren Divergenz- und Leistungsentwicklung

Fach	Effekte	Vortestleistung			Standardabweichung im Vortest		
		df	F	p	df	F	p
Deutsch	Divergenzentwicklung	1/411	27.03	.00	1/411	49.90	.00
	Leistungsentwicklung	1/411	1.88	.17	1/411	13.22	.00
	Divergenz- x Leistungsentwicklung	1/411	1.81	.18	1/411	.45	.50
Englisch	Divergenzentwicklung	1/411	2.08	.15	1/411	36.36	.00
	Leistungsentwicklung	1/411	27.31	.00	1/411	3.35	.07
	Divergenz- x Leistungsentwicklung	1/411	.58	.45	1/411	1.00	.32
Mathematik	Divergenzentwicklung	1/379	.25	.62	1/379	48.38	.00
	Leistungsentwicklung	1/379	5.24	.02	1/379	5.10	.02
	Divergenz- x Leistungsentwicklung	1/379	7.17	.01	1/379	.21	.65

Tabelle 8 : Kennwerte der Diskriminanzfunktion und standardisierte Faktorkoeffizienten

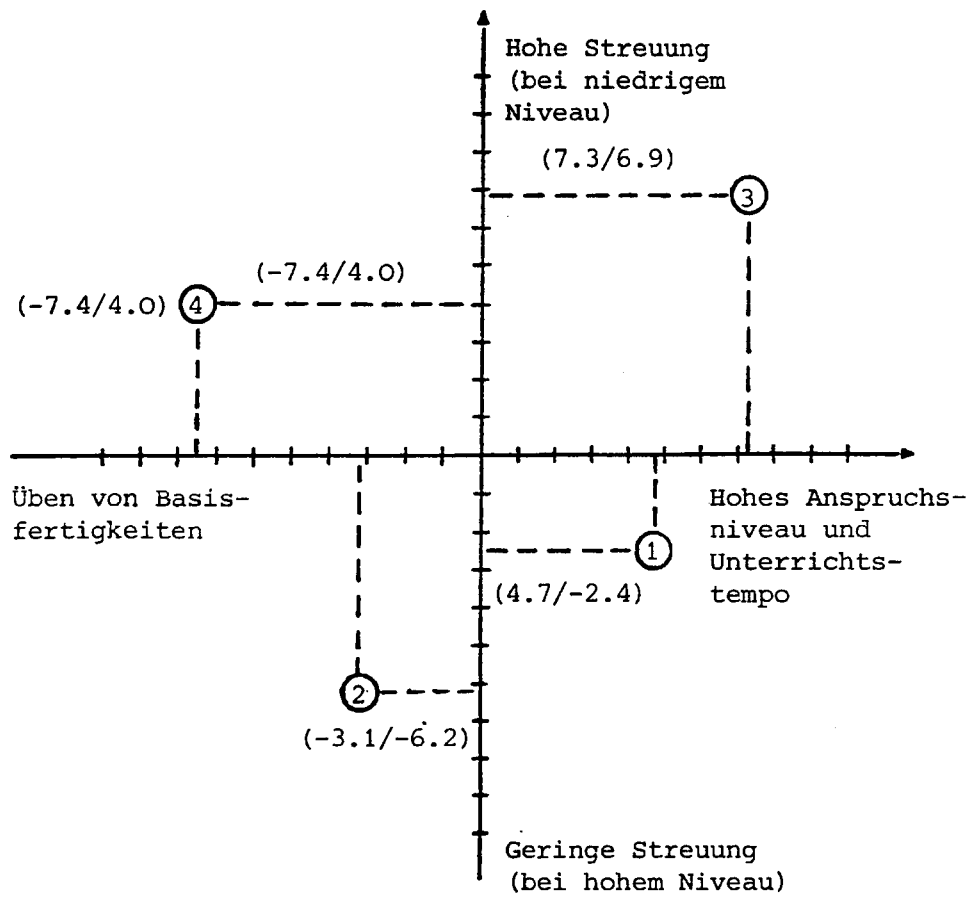
Funktionskennwerte/ diskriminierende Variablen	Koeffizienten
Eigenwert	.23
Kanonische Korrelation	.43
Wilks' Lambda	.81
Vorkenntnisstreuung	.92
Vorkenntnisniveau	-.38
Anspruchsniveau	.28
Stoffmenge	-.66
Übungsintensität	.35
Klarheit der Präsentation	.31
Begründung von Zensuren	-.14

Tabelle 9 : Kennwerte der Diskriminanzfunktionen und standardisierte Faktorkoeffizienten

Funktionskennwerte/ diskriminierende Variablen	Funktion I	Funktion II
Eigenwert	.34	.23
Kanonische Korrelation	.50	.43
Wilks' Lambda	.56	.75
Vorkenntnisniveau	.01	-.47
Vorkenntnisstreuung	.10	.95
Mitarbeit	.03	.13
Anspruchsniveau	.32	.24
Stoffmenge	.45	-.44
Übungsintensität	-.17	.37
Fundamentalstoff	-.59	-.31
Klarheit der Präsentation	.10	.31
Struktur durch Lehrbuch	.21	-.17
Begründung von Zensuren	.02	-.16

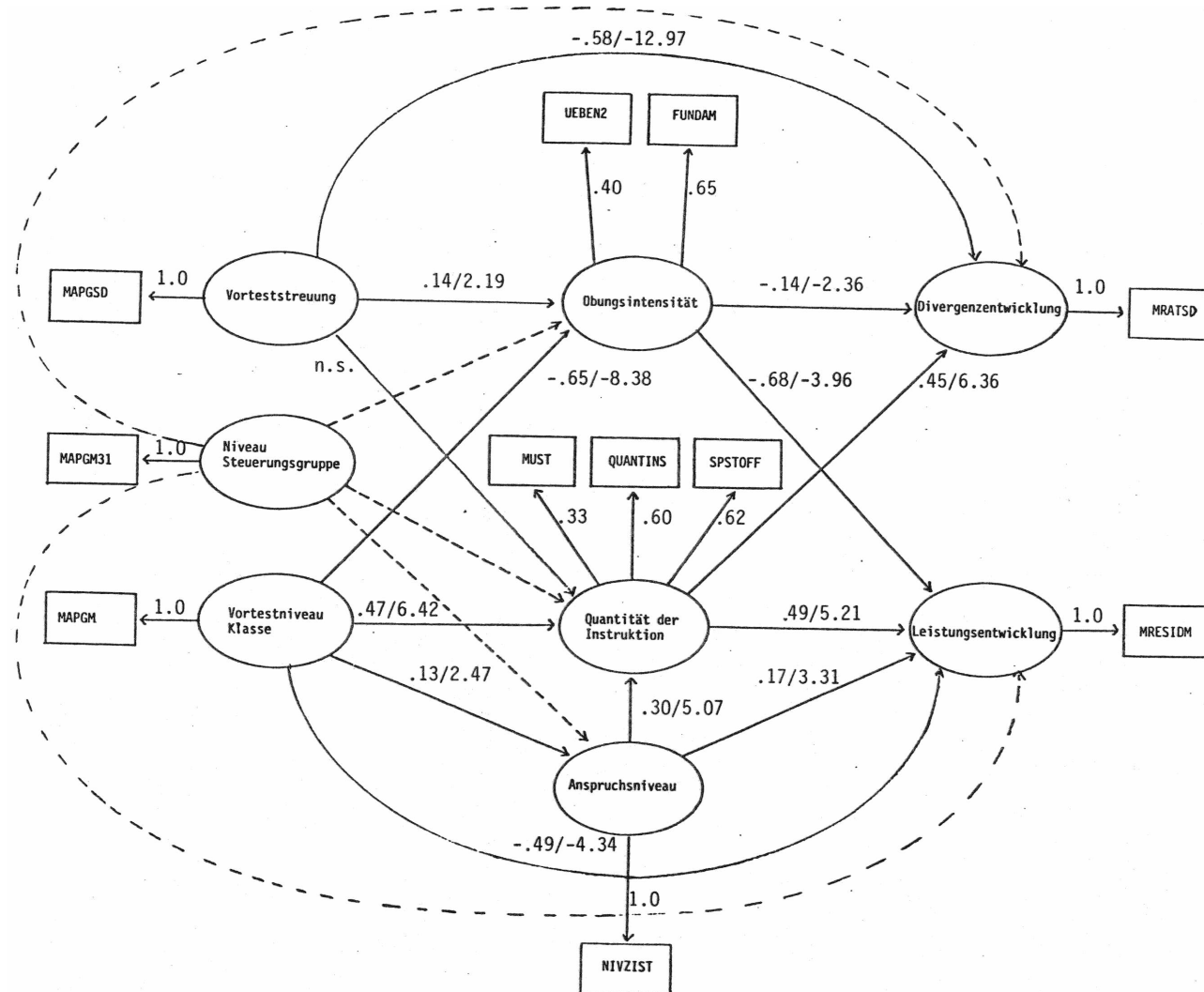


Abbildung 1 : Zentroide der nach Leistungs- und Divergenz-  
entwicklung gruppierten Klassen auf zwei  
Diskriminanzfunktionen<sup>1</sup>



1 Zur Gruppenbildung vgl. Tabelle 6.

Abbildung 2 : Strukturgleichungsmodell zum Zusammenhang von Kontext- und Instruktionsmerkmalen sowie Divergenz- und Leistungsentwicklung (Gesamtmodell und Modellkontrast)



Erläuterungen:

-----> = Koeffizienten im Modell A (im Gesamtmodell alle nicht signifikant)

————> = Koeffizienten im Modell B

.14/2.19 = Koeffizient/t-Wert im Modell B

### 7.3 Skalen und Indizes

Skala/Index	Itemzahl/Indexkonstruktion (Beispiel)	$\alpha$ bzw. M(SD)
Üben von Basisfertigkeiten	9 (häufiges Üben in selbständiger Arbeit während des Unterrichts)	.63
Behandlung von Fundamentalfstoff	6 (Prozentsatz nicht vorausgesetzt)	.82
Anspruchsniveau bei Ziel- und Stoffwahl	6 (wichtiges Ziel ist Strategie der Problemlösung)	.58
Klarheit der Präsentation	4 (Beweise werden vom Lehrer an der Tafel geführt)	.69
Quantität der Instruktion	Prozentsatz im Unterricht behandelte Testitems	51.5(16.1)
Tatsächliche Unterrichtszeit	Anzahl der Unterrichtsstunden zwischen den Meßzeitpunkten	111.3(15.0)

#### 7.4 Literaturverzeichnis

- ACHTENHAGEN, F.: "Mehrebenenanalyse in der Unterrichtsforschung". In: Unterrichtswissenschaft, 9. Jg. (1981), S. 319-336.
- ALDRIDGE, B. G.: "A Mathematical Model for Mastery Learning". In: Journal of Research in Science Teaching, Bd. 20 (1983), H. 1, S. 1-17.
- ALLISON, P. D.: "Measures of Inequality". In: American Sociological Review, Bd. 43 (1978), H. 6, S. 865-880.
- ANDERSON, L., und SCOTT, C.: "The Relationship Among Teaching Methods, Student Characteristics, and Student Involvement in Learning". In: Journal of Teacher Education, Bd. 29 (1978), H. 3, S. 52-57.
- ARLIN, M.: "Teacher Transitions Can Disrupt Time Flow in Classrooms". In: American Educational Research Journal, Bd. 16 (1979), H. 1, S. 42-56.
- ARLIN, M.: "Time Variability in Mastery Learning". In: American Educational Research Journal, Bd. 21 (1984a), H. 1, S. 103-120.
- ARLIN, M.: "Time, Equality, and Mastery Learning". In: Review of Educational Research, Bd. 54 (1984b), H. 1, S. 65-86.
- ARLIN, M., und WEBSTER, J.: "Time Costs of Mastery Learning". In: Journal of Educational Psychology, Bd. 75 (1983), H. 2, S. 187-195.
- ARLIN, M., und WESTBURY, I.: "The Leveling Effect of Teacher Pacing on Science Content Mastery". In: Journal of Research in Science Teaching, Bd. 13 (1976), S. 213-219.
- ATKINSON, R. C.: "Computer-based Instruction and the Learning Process". In: American Psychologist, Bd. 23 (1968), S. 225-239.
- BAKER, D. P., und McPARTLAND, J. M.: Using Longitudinal Test Score Data to Identify Exceptional Learning Environments. Baltimore, Ma.: Johns Hopkins University 1984 (Manuskript).
- BALTES, P. B., u.a.: "On the Dilemma of Regression Effects in Examining Ability Level-related Differentials in Ontogenetic Patterns of Intelligence". In: Developmental Psychology, Bd. 6 (1972), S. 78-84.
- BARR, R.: "How Children are Taught to Read: Grouping and Pacing". In: School Review, Bd. 83 (1975), S. 479-498.

- BARR, R., und DREEBEN R.: "Instruction in Classrooms". In: SHULMAN, L. S. (Hrsg.): Review of Research in Education. Bd. 5, Itasca, Ill. 1977, S. 89-162.
- Bauer, K.-O., u.a.: Sozialisationsbedingungen und soziale Auslese im Sekundarbereich. Dortmund: Arbeitsstelle für Schulentwicklungsforschung (AFS) 1979 (Werkheft Nr. 10).
- BAUER, W.: Lehrerarbeit und Lehrerbewußtsein. München 1980.
- BAUMERT, J.: "Aspekte der Schulorganisation und Schulverwaltung". In: Max-Planck-Institut für Bildungsforschung, Projektgruppe Bildungsbericht (Hrsg.): Bildung in der Bundesrepublik Deutschland. Daten und Analysen. Bd. 1: Entwicklungen seit 1950. Reinbek bei Hamburg und Stuttgart 1980, S. 589-748.
- BAUMERT, J.: "Differenzierung des Sekundarschulangebots oder neue Dreigliedrigkeit?" (erscheint in: Tillmann, K.-J. (Hrsg.): Sekundarschulen unter Konkurrenzdruck. Wiesbaden 1985)
- BAUMERT, J., u.a.: "Chancenausgleich, ein latentes Problem des Gymnasiums?" (erscheint in: Steffens, U. (Hrsg.): Qualität von Schule. Wiesbaden 1986)
- Block, J. H. (Hrsg.): Mastery Learning: Theory and Practice. New York 1971.
- BLOCK, J. H. (Hrsg.): Schools, Society, and Mastery Learning. New York 1974.
- BLOCK, J. H., und BURNS, R. B.: "Mastery Learning". In: SHULMAN, L. S. (Hrsg.): Review of Research in Education. Bd. 4, Itasca, Ill. 1976, S. 3-49.
- BLOOM, B. S.: Human Characteristics and School Learning. New York u.a. 1976.
- BORCHERS, H., und KRÄMER, H.: "Binnendifferenzierung. Erfahrungen im Mathematikunterricht an der Gesamtschule Fröndenberg". In: Gesamtschule, 4. Jg. (1972), H. 1, S. 28-30.
- BORICH, G. D., und FENTON, K. S.: The Appraisal of Teaching. Concepts and Process. Reading, Mass. 1977.
- BROPHY, J. E.: "Teacher Behavior and Its Effects". In: Journal of Educational Psychology, Bd. 71 (1979), H. 6, S. 733-750.
- BROPHY, J. E.: "Successful Teaching Strategies for the Inner-City Child". In: Phi Delta Kappan, Bd. 63 (1982), H. 8, S. 527-530.

- BROWN, W., und SAKS, D. H.: "The Production and Distribution of Cognitive Skills Within Schools". In: Journal of Political Economy, Bd. 83 (1975), S. 571-593.
- BROWN, W., und SAKS, D. H.: An Econometric Perspective on Classroom Reading Instruction. East Lansing: Institute for Research on Teaching, Michigan State University 1978 (Research Series, Nr. 22).
- BURNS, R. B.: "Mastery Learning. Does it Work?" In: Educational Leadership, Bd. 37 (1979), S. 110-113.
- BURSTEIN, L., u.a.: The Use of Within-Group Slopes as Indices of Group Outcomes. Los Angeles: Center for the Study of Evaluation, California University 1981 (CSE Report Nr. 171).
- CARROLL, J. B.: "Ein Modell schulischen Lernens". (1963) In: EDELSTEIN, W., und HOPF, D. (Hrsg.): Bedingungen des Bildungsprozesses. Psychologische und pädagogische Forschungen zum Lehren und Lernen in der Schule. Stuttgart 1973, S. 234-250.
- CRONBACH, L. J., und SNOW, R. E.: Aptitudes and Instructional Methods. New York 1977.
- CRONBACH, L. J., und WEBB, N.: "Between-class and Within-class Effects in a Reported Aptitude and Treatment Interaction. Reanalysis of a Study by G. L. Anderson". In: Journal of Educational Psychology, Bd. 67 (1975), S. 717-724.
- CUSICK, P. A.: Inside High School. The Student's World. New York 1973.
- CUSICK, P. A.: The Egalitarian Ideal and the American High School. New York u.a. 1983.
- DAHLLÖF, U. S.: Ability Grouping, Content Validity and Curriculum Process Analysis. New York 1971.
- DOYLE, W.: "Paradigms for Research on Teacher Effectiveness". In: SHULMAN, L. S. (Hrsg.): Review of Research in Education. Bd. 5, Itasca, Ill. 1977, S. 163-198.
- DOYLE, W.: "Making Managerial Decisions in Classrooms". In: DUKE, D. L. (Hrsg.): Classroom Management. Chicago, Ill. 1979 (Yearbook of the National Society for the Study of Education, Bd. 78,2), S. 42-74.
- EBMEIER, H., und GOOD, T. L.: "The Effects of Instructing Teachers About Good Teaching on the Mathematics Achievement of Fourth Grade Students". In: American Educational Research Journal, Bd. 16 (1979), H. 1, S. 1-16.

- EDELSTEIN, W.: "Das Projekt Schulleistung am Max-Planck-Institut für Bildungsforschung". In: Zeitschrift für Pädagogik, 16. Jg. (1970), S. 517-529.
- EDELSTEIN, W., u.a.: Unterrichtsstoffe und ihre Verwendung in der 7. Klasse der Gymnasien in der BRD. Teil I: Eine empirische Untersuchung. Berlin: Max-Planck-Institut für Bildungsforschung 1968 (Studien und Berichte, Bd. 12).
- EIGLER, G., und STRAKA, G. A.: Mastery Learning. Lernerfolg für jeden? Zielerreichendes Lernen. Erprobung einer herausfordernden Konzeption. München u.a. 1978.
- ENGELHARDT, M. v.: Die pädagogische Arbeit des Lehrers. Eine empirische Einführung. Paderborn u.a. 1982.
- EVERTSON, C., u.a.: "Effects of Class Heterogeneity in Junior High School". In: American Educational Research Journal, Bd. 18 (1981), H. 2, S. 219-232.
- FEND, H.: Gesamtschule im Vergleich. Bilanz der Ergebnisse des Gesamtschulversuchs. Weinheim und Basel 1982.
- FEND, H.: "Determinanten von Schulleistungen. Wie wichtig sind die Lehrer?" In: Unterrichtswissenschaft, 12. Jg. (1984), H. 1, S. 62-67.
- FREDRICK, W. C., und WALBERG, H. J.: "Learning as a Function of Time". In: Journal of Educational Research, Bd. 73 (1980), H. 4, S. 183-194.
- FURBY, L.: "Interpreting Regression Toward the Mean in Developmental Research". In: Developmental Psychology, Bd. 8 (1973), S. 172-179.
- FURCK, C. L.: Das pädagogische Problem der Leistung in der Schule. Weinheim 1961, 5. Aufl., 1975.
- GARNER, W. T.: "The Public Economics of Mastery Learning". In: Educational Technology, Bd. 18 (1978), H. 12, S. 12-17.
- GOOD, T. L.: The Missouri Mathematics Effectiveness Project. A Program of Naturalistic and Experimental Research. Paper presented at the Annual Meeting of the American Educational Research Association, Toronto. Toronto 1978.
- GOOD, T. L.: "Teacher Effectiveness in the Elementary School". In: Journal of Teacher Education, Bd. 30 (1979), H. 2, S. 52-64.

- GOOD, T. L., und POWER, C. N.: "Designing Successful Classroom Environments for Different Types of Students". In: Journal of Curriculum Studies, Bd. 8 (1976), H. 1, S. 45-60.
- GOOD, T. L., u.a.: "Curriculum Pacing: Some Empirical Data in Mathematics". In: Journal of Curriculum Studies, Bd. 10 (1978), H. 1, S. 75-81.
- GUMP, P. V.: The Classroom Behavior Setting. Its Nature and Relation to Student Behavior. Lawrence, Kan.: Psychological Field Station, University of Kansas 1967 (Final Report, Project Nr. 2453).
- HARNISCHFEGER, A., und WILEY, D. E.: "Kernkonzepte des Schullernens". In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, Bd. 9 (1977), H. 3, S. 207-228.
- HAUSSER, K. (Hrsg.): Modelle schulischer Differenzierung. München u.a. 1981.
- HECKHAUSEN, H.: "Chancengleichheit". In: SCHIEFELE, H., und KRAPP, A. (Hrsg.): Handlexikon zur Pädagogischen Psychologie. München 1981, S. 54-60.
- HOPF, D.: Mathematikunterricht. Eine empirische Untersuchung zur Didaktik und Unterrichtsmethode in der 7. Klasse des Gymnasiums. Stuttgart 1980 (Veröffentlichungen aus dem Projekt Schulleistung, Bd. 4).
- HYMAN, J. S., und COHEN, A.: "Learning for Mastery. Ten Conclusions After 15 Years and 3,000 Schools". In: Educational Leadership, Bd. 37 (1979), H. 2, S. 104-109.
- INGENKAMP, K.: Zur Problematik der Jahrgangsklasse. Eine empirische Untersuchung. Weinheim u.a. 1969 (Theorie und Praxis der Schulpsychologie, Bd. 9).
- JACKSON, P. W.: Life in Classrooms. New York 1968.
- JÖRESKOG, K. G., und SÖRBOM, D.: Analysis of Linear Structural Relationships by Maximum Likelihood and Least Squares Methods. Chicago 1981.
- KEESLING, J. W., und WILEY, D.: Regression Models for Hierarchical Data. Paper presented at the Annual Meeting of the Psychometric Society, Stanford University. Stanford 1974.
- KISCHKEL, K.-H.: Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen. Einstellungen, Zufriedenheit und Probleme der Lehrer. Paderborn u.a. 1979 (Arbeitsmaterialien und Berichte zur Sekundarstufe I, Bd. 5).



- KLITGAARD, R. E.: "Going Beyond the Mean in Educational Evaluation". In: Public Policy, Bd. 23 (1975), H. 1, S. 59-79.
- KLITGAARD, R. E., und HALL, G. R.: "A Statistical Search for Unusually Effective Schools". In: FAIRLEY, W. B., und MOSTELLER, F. (Hrsg.): Statistics and Public Policy. Reading, Mass. 1977, S. 51-86.
- KOUNIN, J. S.: Discipline and Group Management in Classrooms. New York u.a. 1970.
- LUNDGREN, U. P.: Frame Factors and the Teaching Process. A Contribution to Curriculum Theory and Theory on Teaching. Diss., Göteborg 1972.
- McPHERSON, A., u.a.: The Effectiveness of Schooling. Edinburgh: University of Edinburgh 1981.
- METZ, M. H.: Classrooms and Corridors. The Crisis of Authority in Desegregated Secondary Schools. Berkeley 1978.
- MORAWIETZ, H.: Flexible Differenzierung im Unterricht. Eine empirische Evaluationsuntersuchung. Oldenburg 1978.
- MORAWIETZ, H.: Unterrichtsdifferenzierung. Ziele, Formen, Beispiele und Forschungsergebnisse. Weinheim und Basel 1980.
- NESSELROADE, J. R., u.a.: "Regression Toward the Mean and the Study of Change". In: Psychological Bulletin, Bd. 88 (1980), H. 3, S. 622-637.
- Petersen, P.: Der kleine Jena-Plan. 1. Aufl., Braunschweig 1927; 56.-60. Aufl., Weinheim und Basel 1980.
- PHILIPP, E., und WITJES, C. W.: "Gymnasium. Abkehr von der Standesschule?" In: ROLFF, H.-G., u.a. (Hrsg.): Jahrbuch der Schulentwicklung. Bd. 2, Weinheim und Basel 1982, S. 123-144.
- RESNICK, L. B.: "Assuming That Everyone Can Learn Everything, Will Some Learn Less?" In: School Review, Bd. 85 (1977), H. 445-452.
- ROEDER, P. M.: Fallstudien zur Fachleistungsdifferenzierung in der Hauptschule. Berlin: Max-Planck-Institut für Bildungsforschung 1984 (Manuskript).
- ROEDER, P. M., u.a.: Sozialstatus und Schulerfolg. Bericht über empirische Untersuchungen. Heidelberg 1965 (Pädagogische Forschungen, Bd. 32).
- ROEDER, P. M., u.a.: "Über Zusammenhänge zwischen Zensur und Testleistung". In: PETILLON, H., u.a. (Hrsg.): Schülergerechte Diagnose. Weinheim (im Druck).

- ROSENSHINE, B.: "Classroom Instruction". In: GAGE, N. L. (Hrsg.): The Psychology of Teaching Methods. Chicago, Ill. 1976 (Yearbook of the National Society for the Study of Education, Bd. 75,1), S. 334-371.
- ROSENSHINE, B. V.: "Content, Time, and Direct Instruction". In: PETERSON, P. L., und WALBERG, H. J. (Hrsg.): Research on Teaching. Concepts, Findings, and Implications. Berkeley, Cal. 1979, S. 28-56.
- SANFORD, J.: "Comparison of Heterogeneous and Homogeneous Junior High Classes. Austin: Research and Development Center for Teacher Education, University of Texas 1980 (R&D Report Nr. 6108).
- SCHEFER, G.: Das Gesellschaftsbild des Gymnasiallehrers. Eine Bewußtseinsanalyse des deutschen Studienrats. Frankfurt a.M. 1969.
- SCHWARTZ, J., und WINSHIP, C.: "The Welfare Approach to Measuring Inequality". In: SCHUESSLER, K. F. (Hrsg.): Sociological Methodology 1980. San Francisco u.a. 1979, S. 1-36.
- SCHWARZER, R., und STEINHAGEN, K. (Hrsg.): Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden. München 1975.
- SIMONS, H., u.a.: "Untersuchungen zur differentialpsychologischen Analyse von Rechenleistungen". In: Zeitschrift für Entwicklungspsychologie, Bd. 7 (1975), H. 3, S. 153-169.
- SMITH, L. M., und GEOFFREY, W.: The Complexities of an Urban Classroom. An Analysis Toward a General Theory of Teaching. New York 1968.
- SÖRENSEN, A. B.: The Organizational Differentiation of Students in School as an Opportunity Structure. Paper presented at the Conference on Conceptualization of School Organization and Schooling Processes, University of Notre Dame, May 13-15, 1985. Notre Dame, Indiana 1985.
- TREIBER, B.: Qualifizierung und Chancenausgleich in Schulklassen. 2 Bde., Frankfurt a.M. 1980 (Europäische Hochschulschriften, Bd. 63 und 64).
- TREIBER, B.: "Classroom Environments and Processes in a Multi-level Framework". In: Studies in Educational Evaluation, Bd. 7 (1981), H. 1, S. 75-84.
- TREIBER, B.: "Lehr- und Lern-Zeiten im Unterricht". In: TREIBER, B., und WEINERT, F. E. (Hrsg.): Lehr-Lern-Forschung. Ein Überblick über Einzeldarstellungen. München u.a. 1982, S. 12-36.

- TREIBER, B., und PETERMANN, F.: "Probleme der Unterrichtsdifferenzierung aus der Sicht des ATI-Forschungsprogramms". In: Zeitschrift für Pädagogik, 22. Jg. (1976), S. 525-546.
- TREIBER, B., und SCHNEIDER, W.: "Qualifizierende und varianzreduzierende Wirkungen von Schulklassenunterricht". In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, Bd. 12 (1980), H. 3, S. 261-283.
- TREIBER, B., und WEINERT, F. E.: Gute Schulleistungen für alle? Psychologische Studien zu einer pädagogischen Hoffnung. Münster 1985 (Arbeiten zur sozialwissenschaftlichen Psychologie, H. 16).
- TREIBER, B., u.a.: "Bedingungen individuellen Unterrichtserfolgs". In: Zeitschrift für Pädagogik, 22. Jg. (1976), H. 2, S. 153-180.
- TREIBER, B., u.a.: "Unterrichtsqualität, Leistungsniveau von Schulklassen und individueller Lernfortschritt". In: Zeitschrift für Pädagogik, 27. Jg. (1982), H. 4, S. 563-576.
- TUKEY, J. W.: Exploratory Data Analysis. Reading, Mass. 1977.
- Viet, U., und Sommer, N.: Entwicklung von Unterrichtseinheiten für den Mathematikunterricht der Gesamtschule und Untersuchung von Differenzierungsmodellen. Abschlußbericht. Osnabrück: Universität Osnabrück, Seminargemeinschaft für Bildungsforschung, Projekt Gesamtschule, Teilprojekt Mathematik 1976 (Manuskript).
- Viet, U., und Tönsing, R.: Untersuchungen über den Einsatz von Lehrprogrammen im Mathematikunterricht bei der Lösung von Differenzierungsproblemen in der Förderstufe (5. und 6. Schuljahr). Abschlußbericht über das von der Stiftung Volkswagenwerk finanzierte Projekt. Osnabrück: Universität Osnabrück 1975.
- WEINERT, F. E., u.a.: "Bedingungsanalysen von (erwartungswidrigen) Schulleistungen". In: HECKHAUSEN, H. (Hrsg.): Fähigkeit und Motivation in erwartungswidriger Schulleistung. Göttingen u.a. 1980 (Motivationsforschung, Bd. 9), S. 106-128.
- WESTBURY, I.: "Conventional Classrooms, Open Classrooms, and the Technology of Teaching". In: Journal of Curriculum Studies, Bd. 5 (1973), S. 99-121.
- WIENOLD, G., u.a.: "Lernmaterial und Lehrerverhalten in institutionalisierten Lehr-, Lern-Prozessen am Beispiel des Englischunterrichts". In: Zeitschrift für Pädagogik, 28. Jg. (1982), H. 4, S. 545-562.

WILEY, D. E.: "Another Hour, Another Day. Quantity of Schooling, a Potent Path for Policy". In: SEWELL, W. H., u.a.: Schooling and Achievement in American Society. New York 1976, S. 225-266.

WOTTAWA, H.: Gesamtschule: Was sie uns wirklich bringt. Eine methodenkritische Darstellung der Schulvergleiche in Hessen, Nordrhein-Westfalen und Niedersachsen. Düsseldorf 1982 (Studien zur Lehrforschung, Bd. 21).

ZEIHER, H.: Gymnasiallehrer und Reformen. Eine empirische Untersuchung über Einstellungen zu Schule und Unterricht. Stuttgart 1973 (Veröffentlichungen aus dem Projekt Schulleistung, Bd. 1).

ZEIHER, H. J.: Unterrichtsstoffe und ihre Verwendung in der 7. Klasse des Gymnasiums in der BRD. Teil II: Deutschunterricht. Berlin: Max-Planck-Institut für Bildungsforschung 1972 (Studien und Berichte, Bd. 24).

REVISION '90

