

Qualifizierende und varianzreduzierende Wirkungen von Schulklassenunterricht

Bernhard Treiber und Wolfgang Schneider

Universität Heidelberg

Schulklassenunterricht soll möglichst vielen Schülern ein möglichst hohes kognitives Leistungsniveau erreichen helfen. Nur wenige Unterrichtsstudien haben bisher jedoch zur Aufklärung jener Lehr-Lern-Bedingungen beigetragen, unter denen sich hohe qualifikatorische und varianzreduktive Wirkungen gleichzeitig erreichen lassen. Die vorliegende Studie greift dafür auf Merkmale einer adaptiv-remedialen Lehr-Lern-Organisation zurück und überprüft ihren Erklärungsgehalt im Mathematikunterricht von 58 fünften Hauptschulklassen. Zur systematischen Entwicklung komplexer Erklärungen und ihrer einzelfallartigen Beurteilung wird ein neu entwickeltes Programmsystem (HYPAG) angewandt. Es führt zur Auswahl einer Annahmenstruktur, in der Instrumental- und Rahmenbedingungen von Schulklassenunterricht in Wechselwirkung stehen, für sich genommen aber nur mehr hinreichende und darin substituierbare Bedingungen qualifizierender und varianzreduktiver Unterrichtseffekte sind.

Schule und Unterricht werden in verschiedenen Bildungsmodellen gewisse *objektive* Funktionen für den Bestand und die Leistungsfähigkeit anderer gesellschaftlicher Teilsysteme (vor allem des Beschäftigungssystems) zugewiesen (vgl. Fend, 1980). Zum Kernbestand einer solchen an Schule und Unterricht von „außen“ herangetragenem Funktionsbestimmung zählt die Anhebung des Leistungsniveaus (= Qualifikation) und der Ausgleich von Leistungsunterschieden (= Varianzreduktion): möglichst viele Schüler eines Jahrgangs sollen ein möglichst hohes gemeinsames Sockelniveau an Fähigkeiten, Fertigkeiten und Kenntnissen erreichen; zugleich sollen bestehende Eingangsunterschiede zwischen Schülern in ihrer Bedeutung für ihr weiteres Lernen reduziert werden. Übereinstimmend wird diesem Zielpaar eine hohe Wertigkeit zugeschrieben, für seine gleichzeitige Realisierung jedoch mit einem Zielkonflikt gerechnet: das Erreichen eines möglichst hohen Qualifizierungsniveaus geht auf Kosten des zugleich angestrebten Ausgleichs von Schülerleistungsunterschieden, der seinerseits nur bei herabgesetzten Niveaueanforderungen erreichbar erscheint (vgl. Heckhausen, 1981). Deshalb wird allenfalls ein Optimum, nicht aber ein Maximum an qualifizierenden *und* varianzreduzierenden Unterrichtseffekten zu

erwarten sein. Dennoch muß im einzelnen geklärt werden, unter welchen Bedingungen sich beide Zielkriterien auch im Schulklassenunterricht harmonisieren lassen oder aber in einen Zielkonflikt münden, der das Erreichen eines der beiden Kriterien nur auf Kosten des jeweils anderen Kriteriums zuläßt.

Auch aus der *subjektiven* Sicht unterrichtender Lehrer nehmen die Ziele der Qualifikation und des Ausgleichs von Leistungsunterschieden eine wichtige handlungsleitende Funktion für ihr eigenes Unterrichtshandeln ein (vgl. v. Engelhardt, 1979; Houston, Roy & Duff, 1972; Houston, Crosswhite & King, 1974; Snow, 1973). So sind Lehrerunterschiede in der subjektiven Wertigkeit dieser Ziele denn auch regelmäßig mit Unterschieden z. B. in der Förderung entweder besonders leistungsfähiger oder leistungsschwacher Schüler verknüpft (vgl. Ritzen, Winkler & Hargreaves-Heap, 1979). Es ist deshalb von theoretischem wie empirischem Interesse, die Auswirkungen eines auf das Erreichen dieser Ziele gerichteten Unterrichts auf kognitive Schülerleistungen zu erfassen.

Vereinzelte Untersuchungen (vgl. Brown & Saks, 1975 a, b; 1978) machen zudem deutlich, daß *einkriteriale* Analyseserien der Lehr-Lern-Forschung die Wirksamkeit von Schule und Unterricht nur unvollständig rekonstruieren können und deshalb z. T. auch zu unzutreffenden Schlußfolgerungen über die durch sie erreichten Wirkungen führen. So kann eine gegebene Schul- oder Unterrichtsvariable mit dem Leistungsniveau einer Schulklasse negativ korreliert sein und dann zu der irrigen Folgerung veranlassen, erhöhte Bemühungen von Lehrern trügen nur zu einer Verringerung des Leistungsniveaus einer Klasse bei. Wird aber auch die — sich möglicherweise gleichfalls verringemde — Leistungsstreuung als zusätzliches Effektkriterium in die „Ertragsbilanzierung“ von Unterricht einbezogen, wird erst offenkundig, daß diese Instrumentalvariablen durchaus wirksam sind — dann aber z. B. eher varianzreduktiv als niveausteigernd. Einkriteriale Lehr-Lern-Analysen würden hingegen leicht auf eine *generelle* Unwirksamkeit von Schule und Unterricht schließen und so den ohnehin stagnierenden Erkenntnisfortschritt in diesem Forschungsbereich (vgl. Brophy, 1979 a, b; Gage, 1979) nur noch verstärken. Dieser Fehlschluß wird aber erst bei *mehrkriterialer* Modellierung der Wirksamkeit von Unterricht nachzuweisen und zu korrigieren sein.

Theoretische, anwendungspraktische wie forschungsstrategische Gründe sprechen demnach dafür, die möglichen Auswirkungen von Schule und Unterricht auf kognitive Schülerleistungen für mehrere Effektkriterien gleichzeitig zu untersuchen und sich dabei vor allem auf qualifikatorische und varianzreduktive Wirkungen zu konzentrieren. Damit werden denn auch jene Bedingungen im herkömmlichen Schulklassenunterricht aufzuklären sein, die ein gleichzeitiges Erreichen eines hohen Qualifizierungsniveaus und einer hohen Varianzreduktion ermöglichen. Die Suche nach dafür geeigneten Erklärungsvariablen bleibt derzeit aber noch durchweg erfolglos: weder finden sich Lehr-Lern-Modelle, die über-

haupt auf die Identifizierung von schulischen Bedingungen mehrkriterialer kognitiver Unterrichtseffekte zielen, noch entsprechende empirische Studien, in denen der Erklärungsgehalt dieser Bedingungsmodelle überprüft werden würde. Dies liegt vor allem daran, daß sich die bisherigen Lehr-Lern-Analysen auf jeweils nur ein Effektkriterium der in Schulklassen erreichten kognitiven Wirkungen richteten, also entweder nur auf das mittlere Leistungsniveau oder die Streubreite bzw. „Leistungsschere“ in Schulklassen. Deshalb müssen, um beide Kriterien gleichzeitig vorhersagen zu können, auch die darauf bezogenen separaten Analyseserien zunächst zusammengelegt werden. Solche integrativen und mehrkriterialen Lehr-Lern-Modelle fehlen derzeit noch. Vorläufer dafür können allenfalls in adaptiv-remedialen Instruktionssystemen (wie IPI, PLAN oder *Mastery Learning*) (vgl. Hambleton, 1974) gesehen werden.

Zu den wichtigsten Merkmalen dieser Systeme gehört ihre — Adaptivität (als Variation des Unterrichtsangebots in Lehrzielen, Lehrmethoden oder Lernzeiten); — Problemsensitivität (als regelmäßige und systematische Diagnose von Lernvoraussetzungen, entstandenen Fähigkeits- und Kenntnislücken, Leistungsschwierigkeiten und vorhandenem Fehlverhalten); — Remedialität (als darauf unmittelbar und spezifisch bezogene didaktische Interventionen des Lehrers) (vgl. dazu ausführlicher Eigler & Straka, 1978; Kleber, Fischer, Hildeschmidt & Lohrig, 1977; Schwarzer, 1977; Treiber, 1980; Weinert, 1977). Übereinstimmend belegen zudem mehrere Übersichtsdarstellungen der empirischen Forschungsergebnisse den qualifizierenden wie varianzreduzierenden Effekt dieser Systemelemente (vgl. Anderson & Block, 1977; Block & Burns, 1976; Block & Anderson, 1975; Bloom, 1976; Ingenkamp, 1979; Barr & Dreeben, 1978; Treiber, 1980). Diese Merkmale können auch handlungstheoretisch (also aus der Sicht des Lehrers) (vgl. Groeben, 1980) oder systemtheoretisch (im Sinne eines ultrastabilen Regelkreises) (vgl. Brandtstädter, 1977) als Komponenten eines flexiblen Problemlösens interpretiert werden: dessen Erfolg ist dann an eben diese Komponenten einer selbstkorrektiven Optimierung mit vielfach effektkontrollierten und ständig rückgekoppelten Such- und Probehandlungen gebunden (vgl. Hunt, 1979; Merrill, 1979; Snow, 1977 b). Doch gibt es bisher noch kaum Studien, in denen diese Elemente einer adaptiv-remedialen Lehr-Lern-Organisation auch zur gleichzeitigen Vorhersage des im herkömmlichen Schulklassenunterrichts erreichten Qualifizierungsniveaus und Varianzausgleichs verwendet wurden. Forschungspraktisch wird deshalb auch zunächst die Entwicklung von Lehr-Lern-Modellen zu fördern sein, in denen adaptiv-remediale Systemmerkmale sich zur Vorhersage dieser beiden Effektkriterien von Unterricht eignen.

Die Entwicklung und Verbesserung solcher Modelle ist jedoch an ein geeignetes prüfmethodisches Instrumentarium gebunden, das die derzeit verfügbaren multivariaten Analyseverfahren nur mehr unzulänglich bereitstellen. Dabei kann auf zumindest die folgenden Mängel hingewiesen werden (Wottawa, 1979):

— So ist es einmal erforderlich, in Erklärungsannahmen zwischen notwendigen und hinreichenden Bildungsbedingungen schulischen Lernens unterscheiden und diese auch prüfmethodisch zutreffend abbilden zu können. Gerade für Lernaufgaben im mittleren Schwierigkeitsbereich ist es nämlich angemessen, in Fähigkeits- und Motivparametern der Schüler, ihrer aufgewandten Lernzeit, der Quantität und Qualität ihres Unterrichts sowie in außerschulischen Lernhilfen jeweils multipel hinreichende Determinanten zu sehen, die sich in verschiedenen Situationen, bei unterschiedlichen Lernanforderungen und für einzelne Personpopulationen teilweise bis vollständig gegenseitig substituieren können (vgl. Heckhausen & Rheinberg, 1980; Krug & Rheinberg, 1980; Weinert, 1980; Weinert & Petermann, 1980; Weinert, Schneider & Treiber, 1980). Die derzeit verfügbaren prüfmethodischen und statistischen Standardversionen multivariater Auswertungsverfahren (z. B. der Kleinst-Quadrat-Schätzungen) stellen dafür aber nur formal einfache Problemlösungen bereit, orientieren sich an linearen und polynominalen Formalstrukturen, die rechentechnisch günstig zu bearbeiten sind, und können so komplexeren Bedingungsannahmen zu schulischem Lernen auch näherungsweise oft nicht mehr genügen (Brown, 1975).

— Die Bestimmung des Geltungsbereichs eines bestimmten Erklärungsmodells wird durch die bisher bevorzugten inferenzstatistischen Prüfverfahren eher erschwert als erleichtert: über die bloße Auskunft hinaus, eine bestimmte (zumeist einfach strukturierte) Hypothese sei mit einem bestimmten zufallskritisch abgesicherten Ergebnismuster verträglich, ist den dabei ermittelten Statistiken nichts zu entnehmen (vgl. Bredenkamp, 1972; Westmeyer, 1979). Insbesondere ist ohne Inspektion des ursprünglichen Datenmaterials in der Stichprobe nicht zu erkennen, auf welche ihrer Elemente (z. B. Schulklassen) sich ein gegebenes Lehr-Lern-Modell tatsächlich bezieht. Der für größere Datenmengen dafür erforderliche Arbeitsaufwand ist aber kaum zu bewältigen und hält denn auch von einer Bestimmung der tatsächlichen Modellgültigkeit an einzelnen Schülern, Schulklassen oder Schulen ab. Vielmehr wird das gerade in Studien zum schulischen Lernen umfangreiche Datenmaterial nur noch in statistischen Kennzahlen zusammengefaßt, die dann auch als empirische Prüfinstanz für theoretische Lehr-Lern-Annahmen genutzt werden. Fast ausschließlich finden sich so nur noch „anonyme“ Erklärungen (vgl. Opp, 1979, S. 17), deren tatsächliches Zutreffen auf einzelne Merkmalsträger und deren empirischer Geltungsbereich folglich offen bleibt.

— Nachteilig erweist sich ferner, daß die dabei verwendeten Kennzahlen (z. B. Mittelwert oder Standardabweichung) bei den oft nur geringen Stichproben der Lehr-Lern-Forschung, deren Analyseeinheiten z. B. vorfindbare Schulklassen oder experimentelle Lerngruppen bilden, für Verteilungseigenschaften und „Ausreißer“ (*outliers*) besonders empfindlich sind (vgl. Andrews, Bickel, Rosers & Tukey, 1972; Barnett & Lewis, 1978; Belsley, Kuh & Welsch, 1979; Launer & Wilkinson, 1979; Wainer, 1976). Deshalb wird neuerdings — vor allem in der sog. Explorativen Datenanalyse (EDA) (vgl. McNeil, 1977; Mosteller & Tukey, 1977; Rey, 1978; Tukey, 1977) — die Verwendung verteilungsrobuster Statistiken (z. B. Mediane oder Interquartilranges) ebenso empfohlen wie eine möglichst einzelfallnahe Bestimmung des tatsächlichen Erklärungsgehaltes verschiedener Modellvarianten und ihrer „interaktiven“ Verbesserung.

— Eine weitere Konsequenz der bisher bevorzugten inferenzstatistischen Prüfverfahren schließlich ist, daß die vorläufige Beibehaltung der statistischen Nullhypothese eines gegebenen Erklärungsmodells weder zu seiner weiteren Elaboration als allgemeines Modell noch zur Ausdifferenzierung verschiedener differentieller oder gar idiografischer Modellvarianten anregt. Gerade beim aktuell erreichten Kenntnisstand und dem künftig benötigten Komplexitätszuwachs der Lehr-Lern-Forschung ist es indes besonders dringlich, ihre bisher stark vereinfachten Erklärungsmodelle schrittweise zu erweitern. Dies gelingt am ehesten durch systematische Verbesserung ihres informativen Gehaltes für einen gegebenen Datensatz, was vor allem die Aufnahme von Zusatzannahmen und Änderungen in der Verknüpfung einzelner Annahmenelemente erforderlich macht. Diesen Vorgang einer interaktiven (theoriegeleiteten und empirisch kontrollierten) Modellentwicklung hat Wottawa (1978, S. 291) als *systematische Hypothesenagglutination* beschrieben und zum Ausgangspunkt einer prüfmethodischen Problemlösung gewählt, die nachfolgend beschrieben wird.

Festzuhalten ist demnach, daß zutreffende mehrkriteriale Erklärungen von Schul- und Unterrichtseffekten derzeit vor allem theoretische Modellentwicklungen voraussetzen, die aber mit den bisher bevorzugten Prüfmethode nur unzulänglich gefördert und z. Zt. sogar verhindert werden. Sie eignen sich insbesondere kaum für die systematische Agglutination von Hypothesen zu komplexeren Modellen und eine explizite Bestimmung ihres Geltungsbereichs am Einzelfall. Inzwischen liegt jedoch ein EDV-Programmsystem (HYPAG) (für Hypothesenagglutination) vor, das auf die Bearbeitung eben dieser Probleme eingeht. Die vorliegende Studie konnte deshalb auch auf diese Analyseverfahren bei der Entwicklung geeigneter Erklärungsmodelle der qualifizierenden und varianzreduzierenden Wirkungen von Schulklassenunterricht zurückgreifen.

In dem im folgenden angewandten Programmsystem HYPAG werden Hypothesen als begriffliche Zusammenfassung von zwei logischen Bewertungsfunktionen dargestellt, die als Zielbedingung (= Dann-Komponente oder Explanandum) und als Vergleichsbedingung (= Wenn-Komponente oder Explanans-Teil) bezeichnet werden (vgl. ausführlicher Wottawa, Härtner & Mattes, 1980). Für eine (meist einfach strukturierte) Zielbedingung wird dann eine Vergleichsbedingung gesucht, die empirisch möglichst oft mit der Zielbedingung übereinstimmt.

Zielbedingung	Vergleichsbedingung	Ergebnisbedingung
wahr	wahr	wahr (erklärt)
wahr	falsch	unerklärt
falsch	wahr	falsch
falsch	falsch	unerklärt

Abb. 1. Mögliche Ergebnisbedingungen bei variabler Verknüpfung der Ziel- und Vergleichsbedingungen als Grundschemata systematischer Hypothesenagglutination (vgl. Wottawa, Härtner & Mattes, 1980)

Ziel der empirischen Hypothesenüberprüfung wird es demnach sein, ein gegebenes Startmodell durch Zusatzhypothesen so zu erweitern und in seiner Annahmenstruktur so abzuändern, daß seine Ziel- und Vergleichsbedingungen möglichst häufig übereinstimmen. Als Bewährungsinstanz wird sich also eine maximale „Trefferquote“ von „wahr-wahr“-Bedingungen eignen, als Falsifikationsinstanz hingegen das (annahmenabhängige) Auftreten einer wahren Vergleichs- und einer falschen Zielbedingung bzw. einer falschen Vergleichs- und einer wahren Zielbedingung. Zur schrittweisen Verbesserung der Beschreibungsgüte eines Modells können diesem dann Zusatzannahmen hinzugefügt, seine Annahmenelemente disjunktiv und konjunktiv zu jeder gewünschten Aussagen zusammengestellt und in ihrem Geltungsbereich durch Intervall- und Schwellenwerte systematisch eingegrenzt werden. Hervorzuheben ist vor allem die Möglichkeit, die Elemente einer theoretischen Bedingungsklasse (z. B. „Unterrichtsvariablen“) untereinander disjunktiv, mit denen einer anderen Klasse (z. B. „Schülerherkunft“) aber konjunktiv (also in geklammerter Form) zu verknüpfen, d. h.

(E_1 oder E_2 oder E_3) und (E_4 oder E_5).

Dabei kann für jeden von n -Fällen das Zutreffen einer Hypothese anhand der resultierenden diskreten Verteilungsfunktion überprüft und auf eine nur inferenzstatistische Beurteilung ihrer spezifischen Beschreibungsgüte somit verzichtet werden. In dieser Studie ist die Zielbedingung (Effektkriterium) das gleichzeitige Auftreten eines hohen Qualifizierungsniveaus und Varianzausgleichs. Die zu ihrer Erklärung geeigneten (unabhängigen) Vergleichsbedingungen sollen dann durch systematische Hypothesenagglutination in Anwendung des HYPAG-Programmsystems schrittweise identifiziert werden.

Die Ziele dieser Untersuchung können demnach so zusammengefaßt werden:

(1) Anders als in bisherigen einkriterialen Lehr-Lern-Analysen geht es in dieser Studie um die Beschreibung und Erklärung eines Zielpaares, nämlich um die qualifizierenden und varianzreduzierenden Wirkungen von Schulklassenunterricht. Beide Kriterien zählen aus subjektiver wie objektiver Sicht zu den Kerngrößen moderner Schulsysteme. Sie werden in dieser Studie über das mittlere Niveau kognitiver Schülerleistungen (= Qualifizierungsaspekt) und deren Entkoppelung von individuellen Schülerunterschieden in ihren Eingangsleistungen (= Varianzreduktionsaspekt) bestimmt.

(2) Zur Erklärung dieses Kriterienpaares werden sog. Metavariablen einer adaptiv-remedialen Lehr-Lern-Organisation von Schulklassenunterricht eingeführt, über die sich Problemlösemerkmale der reflexiven Handlungssteuerung von Lehrern im Unterricht rekonstruieren lassen (vgl. Dörner, 1979; Dörner & Reither, 1978; Reither, 1977, 1979). Vorläufer für diese Erklärungsversion sind individualisierte Lehr-Lern-Systeme (vor allem des zielerreichenden Unterrichts). In dieser Studie sollen Lehrer nach ihrer Unterrichtsorganisation befragt und ihre Aussagen dann in drei Dimensionen („Problemsensitivität“, „Adaptivität“ und „Remedialität“) zusammengefaßt werden.

Die Wirksamkeit dieser Metavariablen kann jedoch nicht unabhängig von vorgegebenen Rahmenbedingungen (vgl. Cooper, Burger & Seymour, 1979; Lundgren, 1972, 1977) bestimmt werden, die das Erreichen eines hohen Qualifizierungsniveaus und Varianzausgleichs mehr oder minder erschweren oder erleichtern können. Dazu zählen üblicherweise institutionelle und organisatorische Normierungen von Schulsystemen, die materielle und personelle Ausstattung einer Schule, aber auch die Größe und Zusammensetzung von Schulklassen nach der Herkunft, Fähigkeit und Lernbereitschaft ihrer Schüler. Dahllöf (1971) untersuchte einige der Implikationen dieser Rahmenbedingungen (*frame factors*) sowohl auf die Organisation des Lehr-Lern-Prozesses als auch auf die individuellen Lernchancen von Schülern. Danach sind mögliche Optionen bei knapper Unterrichtszeit: das Auslassen einzelner Stoffbereiche, die Senkung des Anspruchsniveaus und das Vernachlässigen schwächerer Schüler. Unmittelbar beschränken diese Rahmenbedingungen zumindest die Variation des Unterrichtstempos und damit auch das mittlere erreichbare kognitive Leistungsniveau sowie dessen Streuung innerhalb einer Schulklasse (vgl. Barr & Dreeben, 1978).

In dieser Studie werden jedoch nur zwei Merkmale der begabungsmäßigen Zusammensetzung von Schulklassen herausgegriffen, und zwar das mittlere kognitive Begabungsniveau sowie die Streubreite der Schülerbegabungsunterschiede in einer Klasse. Die Annahme dabei ist, daß ein hohes Begabungsniveau

und eine geringe Begabungsstreuung einer Schulklasse qualifizierende *und* varianzreduzierende Wirkungen leichter erreichen lassen als in weniger günstig zusammengesetzten Vergleichsklassen.

(3) Die logische Verknüpfung dieser Metavariablen untereinander, mit wichtigen Hintergrundvariablen sowie schließlich mit beiden Effektkriterien, die Überprüfung ihres empirischen Erklärungsgehaltes wie auch die Bestimmung ihres Geltungsbereichs stellt an das dafür benötigte prüfmethodische Instrumentarium erhebliche Anforderungen. Dies um so mehr, als Forschungsstand und Bewährungsgrad der gewählten Erklärungsversion für herkömmlichen Schulklassenunterricht noch gering sind. Dabei wird ein neu entwickeltes Programmsystem zur systematischen Hypothesenagglutination (HYPAG) angewandt, das zuvor schon beschrieben wurde.

Die Leithypothesen dieser Studie sind: (1) Schulklassen unterscheiden sich hinsichtlich des in ihnen erreichten Qualifizierungsniveaus und Varianzausgleichs. (2) Unterschiede in diesen Effektkriterien sind auch nach Kontrolle der Zusammensetzung dieser Schulklassen nach Niveauhöhe und Streubreite der allgemeinen kognitiven Fähigkeit ihrer Schüler nachweisbar. (3) Unterschiede in diesen Effektkriterien lassen sich durch Merkmale einer adaptiv-remedialen Lehr-Lern-Organisation erklären.

Der Status dieser Erklärungsvariablen als notwendige und/oder hinreichende Effektbedingungen und ihre Verknüpfungsform untereinander wie auch mit Fähigkeitskovariaten der Schulklassenzusammensetzung ist ungeklärt. Zwar liegt es nahe anzunehmen, daß sich beide Effektkriterien am ehesten bei Wechselwirkung der genannten Instrumental- und Rahmenbedingungen werden erreichen lassen (der ihre konjunktive Verknüpfung entspricht). Dennoch läßt dieses Grundmodell (in Umfang, Struktur und Inhalt) unterschiedliche Varianten zu, einzelne dieser Instrumental- und Rahmenbedingungen als Erklärungselemente aufzunehmen und untereinander zu verknüpfen. Dies macht es notwendig, diese Varianten nacheinander auf ihre Beschreibungsgüte für jede der in dieser Studie untersuchten Schulklassen zu untersuchen.

Methode

Variablen. Die in dieser Untersuchung verwendeten Daten wurden im Rahmen einer größeren Untersuchung zum Mathematikunterricht in 79 fünften Hauptschulklassen im Rhein-Neckar-Kreis erhoben (vgl. Treiber, Weinert & Groeben, 1976; Treiber, in Vorbereitung). Für die folgende Analyse wurden aus diesem Datensatz drei Variablen(gruppen) ausgewählt:

(1) Schülerleistungen in Mathematik: Ein lehrzielorientierter Mathematiktest, dessen 19 Aufgaben den Lehrplananforderungen fünfter Hauptschulklassen entsprachen, nach Angaben der teilnehmenden Lehrer auch im Unterricht behandelt wurden und sich als *Rasch*-skalierbar erwiesen,

wurde zur Schätzung individueller Fähigkeitsparameter von 1943 Schülern zu zwei Meßzeitpunkten im Abstand von jeweils drei Monaten eingesetzt.

(2) Schulklassenzusammensetzung: Die Zusammensetzung der untersuchten Klassen nach Niveau und Streuung ihrer allgemeinen kognitiven Fähigkeiten wurde mit einem Intelligenzsubtest (Wortschatz V1 des Kognitiven Fähigkeitstests KFT 4-13 von Heller, Gaedike & Weinläder, 1976) erfaßt.

(3) Lehr-Lern-Organisation: Zur Beschreibung wichtiger Aspekte einer adaptiv-remedialen Lehr-Lern-Organisation in den untersuchten Klassen wurde ein Fragebogen eingesetzt. Andere Erhebungsmöglichkeiten (wie Interviews oder direkte Unterrichtsbeobachtungen) schieden wegen der Stichprobengröße und aus Rücksichtnahme auf die Belastbarkeit der teilnehmenden Lehrer aus. Die Auswahl geeigneter Fragebogenitems ergab sich in direktem Bezug zu den zuvor genannten Metavariablen („Adaptivität“, „Problemsensitivität“ und „Remedialität“) sowie in Anlehnung an Evaluationsinstrumente der Cooley-Arbeitsgruppe (Cooley & Leinhardt, 1975, 1980) (vgl. auch Ellison & Sherman, 1976; Scheuerman, 1976; Intili, 1978). Die Itemformulierungen beschränkten sich auf die Erhebung der lehrerperzipierten Anwendungshäufigkeit der wichtigsten Elemente einer adaptiv-remedialen Lehr-Lern-Organisation. Nach Experten- und Lehrervorausbefragungen wurden aus der Fragebogenendform drei Skalen zur „Adaptivität“ (mit 30 Items), „Sensitivität“ (10 Items) und „Remedialität“ (20 Items) zusammengestellt, deren Items auf einer vierstufigen Skala mit den Kategorien „häufig-manchmal-selten-nie“ zu beantworten waren (zur Skalenkonstruktion und Itemanalyse vgl. Treiber, 1980).

Durchführung. Im Dezember 1976 bearbeiteten 1943 Schüler aus 79 fünften Hauptschulklassen den Mathematik-Prätest sowie den Intelligenzsubtest. Im März 1977 folgte der Mathematik-Posttest. Dabei machten 58 der teilnehmenden Lehrer unterrichtsbezogene Angaben in einem mehrteiligen Fragebogen, der auch die drei Skalen zur Lehr-Lern-Organisation enthielt.

Ergebnisse

Die individuellen Personparameter der Schüler-Rechenfähigkeit im (Raschskalierten) Vor- und Nachtest wurden für die Indikatorisierung der qualifizierenden und varianzreduzierenden Unterrichtseffekte in den untersuchten Klassen wie folgt verwendet:

Qualifizierungsniveau. Die individuellen Personparameter im Vor- und Nachtest wurden zunächst, zur Bildung von Differenzwerten, voneinander subtrahiert. Die Parameterdifferenzen entsprechen dabei dem individuellen Lernleistungszuwachs, den ein Schüler zwischen den Zeitpunkten t_2 (März 1977) und t_1 (Dezember 1976) auf einem einzigen Kontinuum seiner „Rechenfähigkeit“ erreicht (vgl. Fischer, 1974).

Für die resultierenden Differenzwerte wurden dann die Mediane aller Schulklassen ermittelt, die als Indikatoren des in einem dreimonatigen Unterrichtsintervalls erreichten Qualifizierungszuwachses zu interpretieren sind. Die üblichen Einwände gegen die Verwendung von Differenzwerten (vor allem der Skalenabhängigkeit und Reliabilitätsminderung) sind in dieser Studie, in der sie Niveauänderungen in Schulklassen beschreiben, als hinfällig anzusehen (vgl. Simons & Möbus, 1977, S. 11f.). Aus der Betrachtung der klassenspezifischen

Mediane geht hervor, daß sich die 79 untersuchten Klassen hinsichtlich ihres Qualifizierungszuwachses erheblich unterscheiden.¹⁾

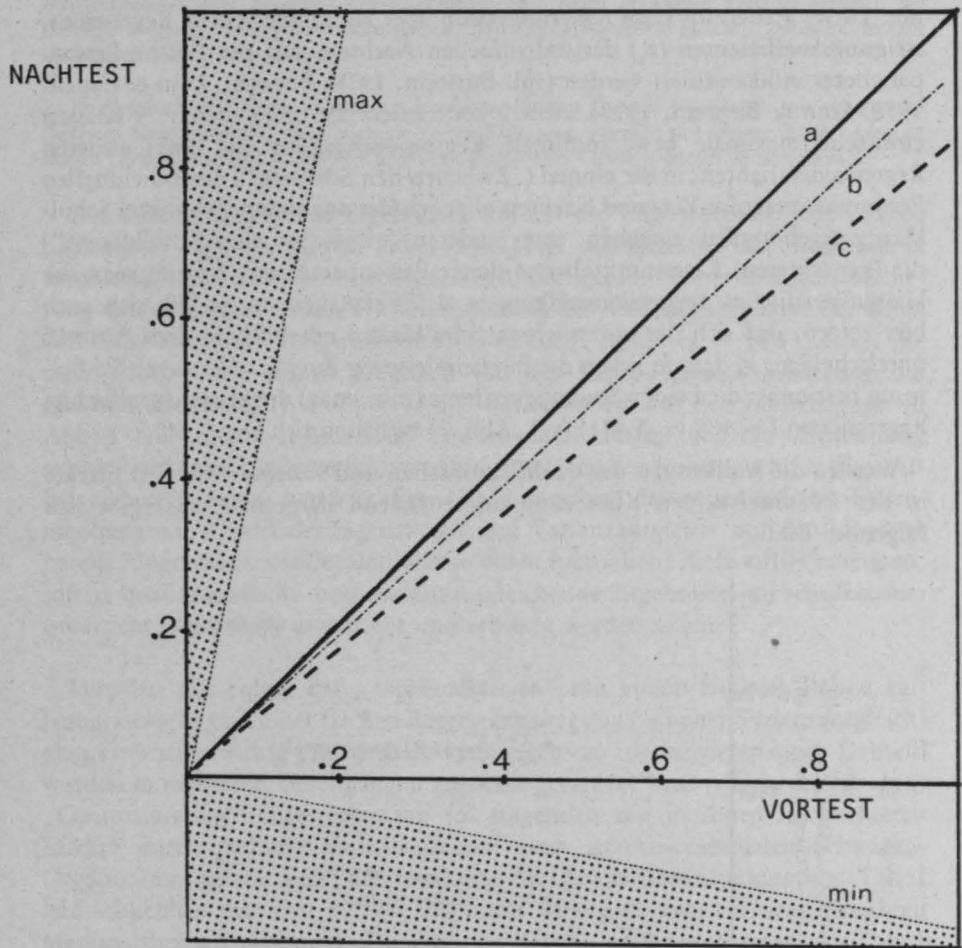


Abb. 2. Verteilungsraum der klassenspezifischen Steigungswinkel
(als Indikator für Varianzreduzierung)

bei Regression der Schülernachtest- auf ihre -vortestleistungen
in 79 fünften Hauptschulklassen mit maximalem, minimalem und mittlerem

- (a) Steigungswinkel sowie den Steigungswinkeln der Regression „Zwischen den Klassen“
(b) und „Zwischen den Schülern“ (c)

1) Ausführliche tabellarische Information hierzu findet sich in einer Langversion des Manuskripts, die bei der Zentralstelle für Psychologische Information und Dokumentation an der Universität Trier, Schneidershof, 5500 Trier, angefordert werden kann.

Varianzreduzierung. Als ein Kriterium für schulischen Leistungsvarianzausgleich wurde die Entkoppelung der individuellen Leistungsvariation zwischen Schülern von Eingangsdifferenzen zu Beginn einer Unterrichtseinheit bezeichnet. Dieses Kriterium kann näherungsweise über klassenspezifische Regressionssteigungskoeffizienten (β_i) der individuellen Nachtest- auf die Vortest-Personparameter indiktorisiert werden (vgl. Burstein, 1978; Burstein, Linn & Capell, 1978; Linn & Burstein, 1977). Abb. 2 kontrastiert die unter diesen 79 Klassen ermittelte maximale bzw. minimale Regressionssteigung mit zwei anderen Regressionsvarianten, in die einmal („Zwischen den Schülern“) die individuellen Personparameter im Vor- und Nachtest aller Schüler unabhängig von ihrer Schulklassenzugehörigkeit eingehen, zum anderen („Zwischen den Schulklassen“) die (gewichteten) Klassenmittelwerte dieser Personparameter. Nimmt man die klassenspezifischen Regressionssteigungen als Effektkriterium, so läßt sich auch hier zeigen, daß sich die untersuchten Schulklassen erheblich in dem Ausmaß unterscheiden, in dem in ihnen die Eingangsleistung ihrer Schüler deren Endleistung bestimmt: die beiden Steigungsextreme (min, max) der klassenspezifischen Regressionen (-0.188 vs. 5.471) (vgl. Abb. 2) verhalten sich wie 1 : 30.

Werden die Indikatoren der qualifikatorischen und varianzreduktiven Effekte in den 79 untersuchten Klassen zusammenfassend dargestellt, so ergibt sich folgendes Bild:

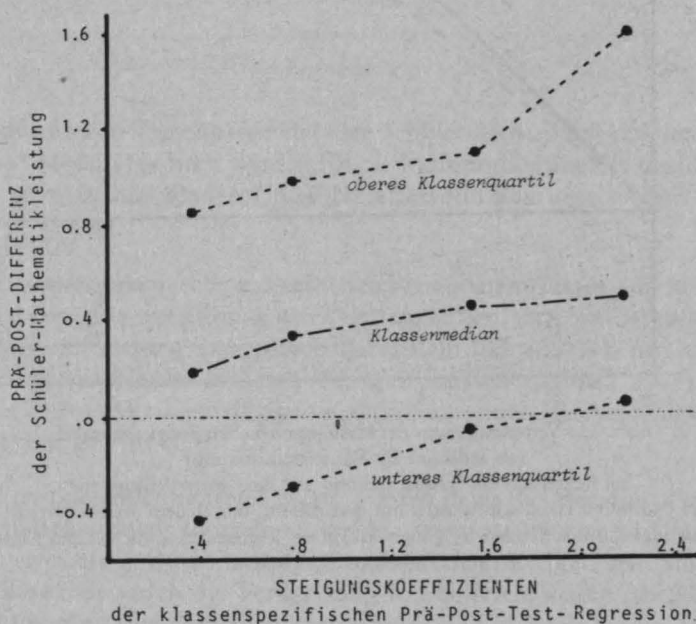


Abb. 3. Grafische Glättung der klassenspezifischen Mediane der Prä-Post-Differenzen der Schülermathematikleistung (als Qualifizierungskriterium für 76 fünfte Hauptschulklassen) sowie des oberen und unteren Klassenquartils bei variablen Steigungskoeffizienten der Regression der Schülerleistung auf ihre Eingangsleistung

Hinsichtlich beider Kriterien unterscheiden sich die hier untersuchten Schulklassen erheblich. Damit kann Hypothese 1 als bestätigt gelten, die diese Effektunterschiede erwartete. Schulklassen mit einem im Durchschnitt höheren Qualifizierungszuwachs ihrer Schüler haben zugleich einen steileren Steigungswinkel ihrer Nachtest-Vortest-Regression (also einen geringeren Varianzausgleich). Bei flacherer Steigung, die einen varianzausgleichenden Schulklasseneffekt indiziert, resultiert ein eher niedrigerer Qualifizierungszuwachs.

Im Vergleich der oberen vs. unteren klassenspezifischen Quartile für Klassen mit zunehmend steilerem Steigungswinkel ihrer Nachtest-Vortest-Regression scheinen sich bessere Schüler (also im oberen Leistungsbereich) gerade in Klassen mit geringerem Varianzausgleich zu verbessern.

Zur Präzisierung dieses Ergebnisses wurden die Indikatorpaare jeder Klasse nach einem Algorithmus von Tukey (1977) (vgl. ausführlicher dazu Leinhardt & Leinhardt, 1980) durch eine „robuste Regressionsgerade“ (über eine Residualisierung der Medianpaare nach Trichotomisierung der Abszisse) approximiert, wobei sich folgende (verteilungsresistente) Gleichung ergab: $y_i = .14 + .22 x_i$. Sie unterscheidet sich einmal erheblich von der Kleinst-Quadrat-Schätzung der gleichen Gleichung ($y_i = .74 + .47 x_i$), was vor allem auf den Einbezug der in Abb. 3 erkennbaren „Ausreißer“ zurückgehen dürfte und die Anwendung explorativer Datenanalysen mit verteilungsrobusten Statistiken auch für diese Arbeit nahelegt. Ferner bestätigt diese Gleichungsform den gegenläufigen Zusammenhang von Qualifizierungszuwachs und Varianzausgleich, auf den eingangs bereits hingewiesen wurde, und läßt so einen (partiellen) Zielkonflikt erkennen, sofern qualifikatorische und varianzausgleichende Ergebnisse im Schulklassenunterricht *gleichzeitig* angestrebt und erbracht werden sollen.

Dies hat zur Folge, daß „Optimalklassen“ mit einem zugleich hohen Leistungszuwachs und einer flachen Regressionssteigung (= hoher Varianzausgleich) eher erwartungswidrig und deshalb auch eigens zu identifizieren sind. Deshalb wurden in mehreren Durchgängen zunächst geeignete Intervallgrenzen für diese „Optimalklassen“ abgesteckt, um im folgenden die in ihnen beobachteten Effekte durch mehrere Modellvarianten einer adaptiv-remedialen Lehr-Lern-Organisation aufzuklären. Die resultierende Quadrantenaufteilung von Tab. 1 ließ schließlich maximal 17 von insgesamt 54 Schulklassen in dem gesuchten Merkmalsbereich erkennen. Die übrigen Klassen verteilen sich, wie in Tab. 1 erkennbar. Dadurch sind zugleich auch die in den untersuchten 54 Klassen gefundenen Kompromißformen klassifiziert, einen (partiellen) Zielkonflikt zwischen dem Erreichen eines hohen Qualifizierungsniveaus vs. Varianzausgleichs aufzulösen:

- zugunsten eher eines hohen Qualifizierungsniveaus auf Kosten von Varianzausgleich (in 21 von 54 Klassen);
- zugunsten von Varianzausgleich auf Kosten des Qualifizierungsaspektes (in 9 von 54 Klassen)

— oder aber, indem in 7 Klassen nur niedrige qualifikatorische und varianz-reduzierende Unterrichtseffekte nachzuweisen sind und so das Auftreten eines Zielkonfliktes ganz umgangen wird.

Tab. 1. Absolute und relative Verteilung von 54 fünften Hauptschulklassen nach der Höhe ihres (dichotomisierten) Qualifizierungsniveaus und Varianzausgleichs

		VARIANZAUSGLEICH			
		hoch	niedrig		
LEISTUNGSZUWACHS	hoch	17 (31.48)	21 (38.89)	38 (70.37)	
	niedrig	9 (16.67)	7 (12.96)	16 (29.63)	
		26 (48.15)	28 (51.85)	54 (100.0)	

Im folgenden wird nun versucht, die in 17 von 54 Schulklassen nachgewiesenen hohen Qualifizierungs- und Varianzreduzierungseffekte durch einzelne Erklärungsvariablen vorherzusagen. Zur Vereinfachung sowohl der aussagenlogischen Modellformulierung wie auch einer einzelfallartigen Inspektion der Modellgültigkeit wurden diese Erklärungsvariablen zunächst dichotomisiert: so wurden zunächst die beiden Kovariaten der Fähigkeitszusammensetzung der untersuchten Klassen (nach Niveauhöhe und Streubreite) anhand der individuellen Testintelligenz ihrer Schüler ermittelt und dann über einen Mediansplit dichotomisiert.²⁾ Ebenso wurden die Fragebogenangaben der Lehrer in den 54 Klassen zur Lehr-Lern-Organisation ihres Unterrichts auf den drei Skalen „Adaptivität“, „Sensitivität“ und „Remedialität“ summiert (was bei ihrer hohen internen Konsistenz zulässig erschien) und über ihren Median auf die Beschreibung einer skalenspezifisch hohen vs. niedrigen Merkmalsausprägung reduziert. Die Menge dieser dichotomisierten Erklärungsvariablen umfaßt somit die beiden Kovariaten von Fähigkeitsniveau und -streuung jeder Klasse und die drei Metamerkmale ihrer Lehr-Lern-Organisation. Die in 17 (von insgesamt 54) Schulklassen beobachteten hohen qualifizierenden und varianzreduzierenden Unterrichtseffekte

²⁾ Tabellarische Information hierzu findet sich in der oben erwähnten Langversion des Manuskripts.

sollten dann durch die folgenden Bedingungen erklärt werden: hohes Fähigkeitsniveau, geringe Fähigkeitsstreuung: Sie gelten als erleichternde Rahmenbedingungen eines leistungseffektiven Schulklassenunterrichts und werden deshalb auch als Kovariaten in ihrer Wirkung auf beide Effektkriterien zu kontrollieren versucht; hohe Adaptivität, Sensitivität und Remedialität der Lehr-Lern-Organisation: Sie werden im folgenden als Instrumentalvariablen von Schulklassenunterricht bezeichnet und sind für schulische Erklärungen der Effekte von Unterricht von besonderem Interesse.

In der vorliegenden Studie gehen diese fünf unabhängigen Erklärungsvariablen also in eine Klasse von Annahmen ein, deren abhängiges Effektkriterium ein hohes Qualifizierungsniveau und ein hoher Ausgleich der Schülerleistungsvarianz ist. Innerhalb dieser Klasse wird zusätzlich eine Wechselwirkung zwischen Instrumental- und Rahmenbedingungen angenommen: es wird also von vornherein unterstellt, daß Merkmale einer adaptiv-remedialen Lehr-Lern-Organisation nur dann wirksam werden, wenn zugleich mindestens eine der beiden Rahmenbedingungen (Fähigkeitsniveau oder -streuung) erleichternd hinzukommt. Deshalb werden Instrumental- und Rahmenbedingungen miteinander auch nur konjunktiv verknüpft. Einzelne Erklärungsvarianten werden sich in dieser Studie also lediglich noch in der Verknüpfung von Instrumental- und Rahmenbedingungen jeweils untereinander unterscheiden, diese also entweder disjunktiv (und damit als hinreichende Bedingungen) oder konjunktiv (und damit als notwendige Bedingungen) in Hypothesenform zusammenstellen. Miteinander werden sie jedoch einheitlich konjunktiv verknüpft, um so ihre postulierte Wechselwirkung auf beide Schulleistungskriterien zutreffend abzubilden.

Dementsprechend wurden verschiedene Modellvarianten durch jeweils unterschiedliche Verknüpfung der genannten Bedingungelemente entwickelt: — und zwar als *reduzierte* Modelle, in die einmal nur die beiden Kovariaten bzw. die drei Instrumentalvariablen eingingen; — als *vollständige* Modelle, in denen die beiden Kovariaten mit den übrigen drei Instrumentalvariablen verknüpft werden. Dies sollte einen Vergleich der durch schulische vs. außerschulische Bildungsbedingungen erreichten Effekterklärung ermöglichen.

Abb. 4 gibt diese Modellvarianten wieder und ordnet ihnen zugleich die Ergebnisse ihrer Überprüfung durch das HYPAG-Programm zu. Die empirische Prüfinstanz für ihre Beschreibungsgüte ist dabei die Anzahl der „wahr-wahr“-Fälle (von maximal 17 Schulklassen), in denen Ziel- und Vergleichsbedingungen vollständig übereinstimmen und in denen sich beide Effektkriterien durch eine spezifische Verknüpfung von Rahmen- und Instrumentalbedingungen von Schulklassenunterricht zutreffend aufklären lassen. Die Ergebnisse zeigen, daß reduzierte Modellvarianten, in die entweder nur die beiden Kovariaten der begabungsmäßigen Klassenzusammensetzung (Modell 1 und 2) eingehen oder in denen lediglich Instrumentalvariablen der Lehr-Lern-Organisation verknüpft werden (Modell 3 und 4), gerade in ihrer disjunktiven Form zwar auf hinreichend viele der 17 „optimalen“ Schulklassen zutreffen: nämlich auf 13 Klassen in Modell 2 bzw. auf 14 Fälle in Modell 4.

MODELL	WENN-KOMPONENTE (= Vergleichsbedingung)				DANN-KOMPONENTE (= Zielbedingung)			ERGEBNISBEDINGUNG			
	Instrumentalbedingungen		Rahmenbedingungen		hohes Qualifi- zierungs- niveau	hohe Varianz- reduzierung	erklärt	unerklärt	falsch	unerklärt	
	hohe Adapti- vität	hohe Remedi- alität	hohe Sensiti- vität	hohes Fähig- keits- niveau							geringe Fähig- keits- steuerung
1 reduziert					und		und	8	9	29	8
2 reduziert					oder		und	13	4	14	23
3 reduziert		und	und				und	0	17	35	2
4 reduziert		oder	oder				und	14	3	16	21
5 vollständig		oder	oder	und	und		und	11	6	16	21
6 vollständig		oder	oder	und	oder		und	12	5	19	18

Abb. 4. Vergleich der Verknüpfungsmöglichkeiten und Beschreibungsgüte verschiedener Modellelemente zur Aufklärung eines hohen Qualifizierungsniveaus und Varianzausgleichs für die Schülermathematikleistungen in 54 fünften Hauptschulklassen durch systematische Hypothesenagglutination mit dem HYPAG-Programmsystem.

Werden Rahmen- und Instrumentalbedingungen aber miteinander verknüpft (Modell 5 und 6), so zeigt sich, daß die reduzierten Modellvarianten (1—4) in 11 Fällen auf die jeweils gleichen Klassen zutrafen. Rahmen- und Instrumentalbedingungen sind demnach im wesentlichen keine spezifischen Erklärungsfaktoren, sondern stehen in enger Wechselwirkung. Werden sie folglich miteinander konjunktiv (und untereinander disjunktiv) verknüpft, so erklärt diese Modellvariante die Unterrichtseffekte von immerhin 12 (von maximal 17) Schulklassen. Ihre beiden notwendigen Bedingungsklassen werden dabei durch nur hinreichende (weil gegenseitig substituierbare) Instrumentalfaktoren aufgefüllt: wahlweise trägt eines der drei lehrlern-organisatorischen Metamerkmale zur Erbringung der erwünschten Unterrichtseffekte bei, wenn zugleich ein hohes Fähigkeitsniveau oder eine geringe Fähigkeitsstreuung einer Schulklasse für günstige Rahmenbedingungen sorgt. Diese aussagenlogische Verknüpfung der Bedingungelemente in der Wenn-Komponente dieser „optimalen“ Modellvariante faßt nochmals Abb. 5 zusammen.

Diskussion

Die vorliegenden Ergebnisse besagen zunächst folgendes: das gleichzeitige Erreichen eines hohen Qualifizierungsniveaus und ein erfolgreiches Entkoppeln von Schülerleistungsunterschieden von bestehenden Eingangsunterschieden kann weder durch Rahmenbedingungen der begabungsmäßigen Zusammensetzung dieser Klassen allein noch durch Merkmale ihrer Lehr-Lern-Organisation hinreichend aufgeklärt werden. Beide Bedingungsklassen stehen vielmehr in Wechselwirkung. Zwar schränken die zu Beginn des Untersuchungsintervalls erhobenen Rahmenbedingungen den nachfolgend nutzbaren Spielraum für eine qualifizierende und varianzausgleichende Beeinflussung von Schulleistungen bereits ein (Modell 2); ebenso können die lehr-lern-organisatorischen Instrumentalvariablen erfolgreich mit optimalen Unterrichtseffekten zumindest in mehreren Klassen in Verbindung gebracht werden (Modell 4); dennoch gelingt es erst einer konjunktiven Verknüpfung beider Bedingungsklassen (Modell 6), die Unterrichtseffekte in 12 von 17 Optimalklassen aufzuklären, ohne dabei die relative Bedeutung einer der beiden Bedingungsklassen zu überschätzen. Dies ist zugleich ein wichtiger empirischer Beleg dafür, daß sich auch mehrkriteriale Unterrichtseffekte theoretisch gehaltvoll und empirisch erfolgreich aufklären lassen.

Einzuräumen ist zwar, daß die Erklärungsbedingungen dieser Unterrichtseffekte mit Erhebungsverfahren beschrieben wurden, die verschiedene Mängel aufweisen. Dies gilt insbesondere für den Lehrerfragebogen zur Erfassung der lehrerperzipierten Anwendungshäufigkeit wichtiger Elemente einer adaptiv-remedialen Lehr-Lern-Organisation. Diese Auskunftsource ist nachweislich nicht immer hinreichend zuverlässig (vgl. Hook & Rosenshine, 1979), und im

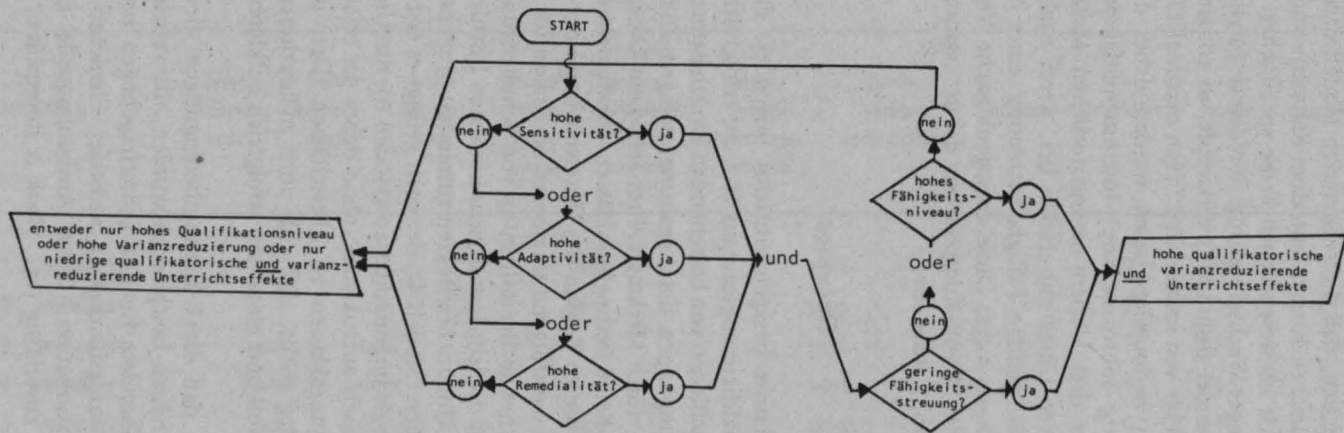


Abb. 5. Aussagenlogische Verknüpfung der Elemente des Optimalmodells zur Aufklärung des hohen Qualifizierungsniveaus und Varianzausgleichs in 54 fünften Hauptschulklassen

Vergleich von Lehrerangaben in verschiedenen Fragebögen zum gleichen Merkmalsbereich wie im Vergleich verschiedener Datenquellen (durch Lehrer- oder Schülerbefragung oder durch Fremdbeobachtung) zur Verankerung der gleichen Unterrichtsmerkmale ergeben sich denn auch häufig erhebliche Diskrepanzen (vgl. Seidman, Linney & Rappaport, 1979). Die in dieser Studie ermittelten empirischen Befunde, die Identifikation der 17 „Optimalklassen“ und die Erklärung ihrer Unterrichtseffekte durch lehr-lern-organisatorische Unterrichtsmerkmale sind denn auch mehr oder minder methodenabhängig; andere Beschreibungsinstrumente und Datenquellen können dann ebenso zu anderen Erklärungsmodellen führen und diese erfolgreich bestätigen wie sie die hier gewählte Modellform in Frage stellen mögen.

Die in dieser Studie ermittelten Ergebnisse besagen jedenfalls, daß nur bei günstigen Rahmenbedingungen eine adaptiv-remediale Lehr-Lern-Organisation zu einem hohen Qualifizierungsniveau und Varianzausgleich führt, und daß andererseits die Wirksamkeit der begabungsmäßigen Klassenzusammensetzung als Rahmenbedingung auf die Unterstützung sie begünstigender Unterrichtsbedingungen angewiesen ist. So erleichtert ein hohes Begabungsniveau oder eine geringe Begabungsstreuung zu Beginn einer Unterrichtseinheit zwar, daß — quasi „naturwüchsig“ — auch die an ihrem Ende ablesbaren Schülerleistungen ein höheres Niveau erreichen und von ihrer Eingangsleistung weniger abhängig sind als in anders zusammengesetzten Vergleichsklassen. Entscheidend ist jedoch, daß diese lernerleichternde Rahmenkonstellation erst durch das Wirksamwerden lehr-lern-organisatorischer Unterrichtsvariablen voll bedeutsam wird. Ähnliche Schüler-Unterrichts-Wechselwirkungen wurden bisher nur auf der Individualebene (vgl. Cronbach & Snow, 1977; Snow, 1977a) nachgewiesen; sie lassen sich aber offenbar auch auf der Aggregatebene von Schulklassen zur Vorhersage ihrer Unterrichtseffekte sichern, obwohl dabei mit z. T. unterschiedlichen Effektmustern zu rechnen ist (vgl. Corno, 1979; Greene, 1979). Sie korrigieren damit die Ergebnisse von Untersuchungen, in denen die Unterrichtseffekte von Merkmalen adaptiv-remedialer Lehr-Lern-Organisation unter innovativen Schul- oder experimentellen Unterrichtsbedingungen als Haupteffekte modelliert werden konnten. Dieses Wirkmuster läßt sich unter den Bedingungen herkömmlichen Schulklassenunterrichts aber nicht bestätigen: hier kommt es offenbar wesentlich auf die gleichzeitige Unterstützung begünstigender Rahmenbedingungen an, die einzelnen Unterrichtsvariablen vorgegeben sind. Dazu zählt auch die begabungsmäßige Zusammensetzung einer Klasse, ohne die oder gar gegen die auch eine an sich effektive Lehr-Lern-Organisation hohe qualifizierende und varianzreduzierende Unterrichtseffekte nicht gleichzeitig erbringen kann. Sie bleibt dann vielmehr hinter ihrem Zielkriterium zurück: entweder wird nur eines dieser Effektkriterium auf Kosten des jeweils anderen erzielt, oder beide in nur reduzierter Form.

Eine weitere Restriktion, denen die Wirksamkeit adaptiv-remedialer Instrumentalvariablen im herkömmlichen Schulklassenunterricht unterliegt, läßt sich aus dem Vergleich der Annahmestruktur der beiden reduzierten Modelle 3 und 4 (in Abb. 4) erschließen: danach erweist sich nur eine disjunktive Verknüpfung der drei Instrumentalvariablen als hinreichend erklärungs-fähig, nicht hingegen eine konjunktive Verknüpfungsform. Wie die Ergebnisse unter idealisierten Untersuchungsbedingungen (also aus Experimental- und Trainingsstudien) zeigen, tragen sämtliche der hier erfaßten Organisationsmerkmale (Adaptivität, Sensitivität und Remedialität) wesentlich zur Aufklärung kognitiver wie affektiver Unterrichtseffekte bei. Davon ist zunächst auch im vorliegenden Falle auszugehen. Daß diese Merkmale dennoch nicht gemeinsam (also bei konjunktiver Verknüpfung) mit beiden Effektkriterien erfolgreich in Verbindung gebracht werden konnten, deutet denn auch eher auf Restriktionen in ihrer gleichzeitigen Durchsetzung unter herkömmlichen Lehr-Lern-Bedingungen hin, als daß ihre wahlweise Nutzung bereits zur Erbringung der erfaßten Bildungsleistungen ausreichte. Zwar muß dabei eine gewisse Überlappung (Multikollinearität) dieser Instrumentalvariablen untereinander in Rechnung gestellt werden. Dennoch ist zusätzlich anzunehmen, daß sich die Integration der erfaßten lehr-lern-organisatorischen Merkmale in einem adaptiv-remedialen Unterrichtsmodell unter den vorfindbaren Schulbedingungen zu voraussetzungsvoll wie auch als zu aufwendig erweist: zu voraussetzungsvoll, weil sie an die Lehrkompetenz, die Sach- und Materialausstattung in einer Klasse sowie an die Unterstützung durch die Schüler (und deren Eltern) erhebliche (und möglicherweise unrealistisch hohe) Anforderungen stellt. Aber auch zu aufwendig, weil die Nutzung von wahlweise nur einem dieser Organisationsmerkmale zur Erbringung der in diesen Klassen erfaßten Unterrichtseffekte (unter zugleich günstigen Rahmenbedingungen) ausreichte. Demnach wäre zu erwarten, daß sich unter veränderten Schulbedingungen, die der Unterrichtsgestaltung in Schulklassen einen größeren Spielraum offen halten, mit einer größeren Varianz zwischen den Schulklassen sowohl auf der Effekt- wie auf der Bedingungsseite zu rechnen wäre, die dann auch eine konjunktive Verknüpfung der kritischen Instrumentalvariablen für die Vorhersage anspruchsvollerer Unterrichtseffekte als den hier nachweisbaren erforderten.

Die in dieser Studie gewählte Form der Datenanalyse ermöglichte es jedenfalls, für jede der 17 „Optimalklassen“ im einzelnen nachzuprüfen, welche Modellstruktur am ehesten mit ihrer Bedingungs- und Effektkonstellation verträglich ist. Die in Abb. 4 als Modell 6 aufgeführte Annahmestruktur gibt für die Erklärung der Unterrichtseffekte in diesen „Optimalklassen“ insofern nur den allgemeinen Rahmen ab, der durch seine nur disjunktive Verknüpfungsform einer weiteren Modellentwicklung z. B. für umschriebene Teilgruppen von Schulklassen oder gar einzelne Klassen hinreichend Raum läßt. Er führt die kognitiven Wirkungen von Unterricht nämlich auf nur mehr hinreichende Bedingungen

in Schule und Unterricht zurück, die deshalb auch (teil)substituierbar und notwendig nur an für sie günstige Rahmenbedingungen gebunden sind. Daraus resultiert jene Elastizität der Determinanten individueller wie aggregierter Bildungsleistungen, auf die neuere Untersuchungen zunehmend häufiger eingehen (vgl. Treiber, 1980). So wird denn auch in Folgestudien die hier ermittelte allgemeine Annahmenstruktur disjunktiv verknüpfter Instrumentalvariablen in Wechselwirkung mit Rahmenbedingungen der begabungsmäßigen Klassenzusammensetzung zunächst repliziert werden müssen, um die bisher allzu schmale empirische Datenbasis für mehrkriteriale Erklärungsmodelle zu erweitern und deren Geltungsbereich kritisch abzuschätzen.

Summary

Classroom teaching is meant to help as many students as possible to reach a maximum cognitive mastery level. There are only a few empirical classroom studies, however, that have contributed to explicate the teaching-learning conditions of high qualificatory and equalizing achievement effects. For that, components of an adaptive-remedial teaching-learning organization are used in the present study and evaluated empirically for mathematics in 58 fifth grades. A newly developed computer program (HYPAG) is also employed which allows to systematically develop complex explanations and test them empirically for each individual event case. The results show an interaction of instrumental and frame conditions of classroom teaching, which are themselves only sufficient (and therefore to-be-substituted) conditions of qualificatory and equalizing achievement effects.

Literatur

- Anderson, L. W. & Block, J. H., Mastery learning. In D. Treffinger, J. Davis & R. Ripple (Hg.), Handbook of educational psychology: Instructional practice and research. New York: Academic Press, 1977. S. 163—186.
- Andrews, D. F., Bickel, P. J., Rosers, W. H. & Tukey, J. W., Robust estimates of location. Princeton, N. J.: Princeton University Press, 1972.
- Barnett, V. & Lewis, T., Outliers in statistical data. New York: Wiley, 1978.
- Barr, R. & Dreeben, R., Instruction in classrooms. In L. S. Shulman (Ed.), Review of research in education. Vol. 5. Itasca, Ill.: Peacock, 1977. S. 89—162.
- Belsley, D. A., Kuh, E. & Welsch, R. E., Regression diagnostics: Identifying disparate data and sources of collinearity. New York: Wiley, 1979.
- Block, J. H. & Anderson, L. W., Mastery learning in classroom instruction. New York: MacMillan, 1975.
- Block, J. H. & Burns, R. B., Mastery learning. In L. S. Shulman (Hg.), Review of research in education. Vol. 4. Itasca, Ill.: Peacock, 1976. S. 3—49.
- Bloom, B. S., Human characteristics and school learning. New York: McGraw-Hill, 1976.

- Brandtstädter, J., Gedanken zu einem psychologischen Modell optimaler Entwicklung. In J. Schneider & M. Schneider-Düker (Hg.), *Interpretationen der Wirklichkeit*. Ernst E. Boesch zum 60. Geburtstag. Saarbrücken: ssip-Schriften, 1977.
- Bredenkamp, J., *Der Signifikanztest in der psychologischen Forschung*. Frankfurt: Akademische Verlagsanstalt, 1972.
- Brophy, J. E., Teacher behavior and its effects. *Journal of Educational Psychology*, 1979, 71, 733—750 (a).
- Brophy, J. E., *Advances in teacher effectiveness research*. Occasional Paper No. 18. East Lansing: Institute for Research on Teaching. Michigan State University, 1979 (b).
- Brown, D. J., Mirror, mirror . . . Down with the linear model. *American Educational Research Journal*, 1975, 12, 491—505.
- Brown, W. & Saks, D. H., The production and distribution of cognitive skills within schools. *Journal of Political Economy*, 1975, 83, 571—593 (a).
- Brown, W. & Saks, D. H., Proper data aggregation for economic analysis of school effectiveness. *Public Data Use*, 1975, 3, 13—18 (b).
- Brown, W. & Saks, D. H., An econometric perspective on classroom reading instruction. Research Series No. 22. East Lansing: Institute for Research on Teaching, Michigan State University, 1978.
- Burstein, L., Assessing differences between grouped and individual-level regression coefficients: Alternative approaches. *Sociological Methods and Research*, 1978, 7, 5—28.
- Burstein, L., Linn, R. L. & Capell, F. J., Analyzing multilevel data: The presence of heterogeneous slopes within class regressions. *Journal of Educational Statistics*, 1978, 3, 347—383.
- Cooley, W. W. & Leinhardt, G., *Design for the Individualized Instruction Study: Study of the effectiveness of individualized instruction in the teaching of reading and mathematics in compensatory education programs: Final report*. Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1975.
- Cooley, W. W. & Leinhardt, G., *The Instructional Dimensions Study*. To appear in: *Educational Evaluation and Policy Analysis*, 1980.
- Cooper, H. M., Burger, J. M. & Seymour, G. E., Classroom context and student ability as influences on teacher perceptions of classroom control. *American Educational Research Journal*, 1979, 16, 189—196.
- Corno, L., A hierarchical analysis of selected naturally occurring aptitude-treatment interactions in the third grade. *American Educational Research Journal*, 1979, 16, 391—410.
- Cronbach, L. J. & Snow, R. E., *Aptitudes and instructional methods*. New York: Irvington, 1977.
- Dahllöf, U., *Ability grouping, content validity, and curriculum process analysis*. New York: Teachers College Press, 1971.
- Dörner, D., Kognitive Merkmale erfolgreicher und erfolgloser Problemlöser beim Umgang mit sehr komplexen Systemen. In H. Ueckert & D. Rhenius (Hg.), *Komplexe menschliche Informationsverarbeitung*. Stuttgart: Huber, 1979. S. 185—195.
- Dörner, D. & Reither, F., Über das Problemlösen in sehr komplexen Realitätsbereichen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 1978, 25, 527—551.

- Eigler, G. & Straka, G. H., *Mastery Learning: Lernerfolg für jeden?* München: Urban & Schwarzenberg, 1978.
- Ellison, J. L. & Sherman, T. M., *Systematic process for evaluating the individualization of instruction.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1976.
- Engelhardt, M. v., *Qualifikation und Selektion in der Schule — Pädagogische Arbeitsorientierungen und gesellschaftliches Bewußtsein von Lehrern.* Zeitschrift für Soziologie, 1979, 8, 111—128.
- Fend, H., *Theorie der Schule.* München: Urban & Schwarzenberg, 1980.
- Fischer, G. H., *Einführung in die Theorie psychologischer Tests.* Bern: Huber, 1974.
- Gage, N. L., *Unterrichten — Kunst oder Wissenschaft?* München: Urban & Schwarzenberg, 1979.
- Greene, J. C., *Class effects in ATI's.* Paper presented at the Annual Meeting of the American Educational Research, San Francisco, 1979.
- Groeben, N., *Die Handlungsperspektive als Theorierahmen für Forschung im pädagogischen Feld.* In M. Hofer (Hg.), *Informationsverarbeitung und Entscheidungsverhalten von Lehrern.* München: Urban & Schwarzenberg, 1980.
- Hambleton, R. K., *Testing and decision-making procedures for selected individualized instructional program.* Review of Educational Research, 1974, 44, 371—400.
- Heckhausen, H., *Stichwort: Chancengleichheit.* In H. Schiefele & A. Krapp (Hg.), *Handlexikon Pädagogische Psychologie.* München: Ehrenwirth, 1981.
- Heckhausen, H. & Rheinberg, F., *Lernmotivation im Unterricht, erneut betrachtet.* Unterrichtswissenschaft, 1980, 8, 7—47.
- Heller, M. S., Gaedike, A.-K. & Weinläder, H., *Kognitiver Fähigkeitstest für 4. bis 13. Klassen (KFT 4—13).* Weinheim: Beltz, 1976.
- Hook, C. M. & Rosenshine, B. V., *Accuracy of teacher reports of their classroom behaviors.* Review of Educational Research, 1979, 49, 1—12.
- Houston, S. R., Roy, M. R. & Duff, W. L., *Judgment analyses as a technique for evaluating school effectiveness.* Journal of Experimental Education, 1972, 40, 56—61.
- Houston, S. R., Crosswhite, C. E. & King, R. S., *The use of judgment analysis in capturing student policies of rated teacher effectiveness.* Journal of Experimental Education, 1974, 43, 28—34.
- Hunt, D. E., *The new three R's in person-environment interaction: responsiveness, reciprocity, and reflexivity.* Tijdschrift voor Onderwijsresearch, 1979, 4, 184—190.
- Ingenkamp, F.-D., *Zielreichendes Lernen — Mastery Learning.* Ravensburg: Otto Maier Verlag, 1979.
- Intili, J. A., *Use of resources in individualized classrooms.* Paper presented at the Annual Meeting of the American Educational Research Association, Toronto, 1978.
- Kleber, E. W., Fischer, R., Hildeschiedt, A. & Lohrig, K., *Lernvoraussetzungen und Unterricht. Zur Begründung und Praxis adaptiven Unterrichts.* Weinheim: Beltz, 1977.
- Krug, S. & Rheinberg, F., *Erwartungswidrige Schulleistungen im Entwicklungsverlauf und ihre Ursachen: Ein überholtes Konstrukt in neuer Sicht.* In H. Heckhausen (Hg.), *Fähigkeit und Motivation in erwartungswidriger Schulleistung.* Göttingen: Hogrefe, 1980.

- Launer, R. L. & Wilkinson, G. N. (Hg.), *Robustness in statistics*. New York: Academic Press, 1979.
- Leinhardt, G. & Leinhardt, S., *Exploratory data analysis: new tools for the analysis of empirical data*. In D. Berliner (Hg.), *Review of Research in Education*. Vol. 8. American Educational Research Association, 1980.
- Linn, R. C. & Burstein, L., *Descriptors of aggregates*. Report Series, Center for the Study of Evaluation. University of California, Los Angeles, 1977.
- Lundgren, U. P., *Frame factors and the teaching process*. Stockholm: Almqvist & Wiksell, 1972.
- Lundgren, U. P., *Model analysis of pedagogical processes*. Stockholm: CWK Gleerup, 1977.
- McNell, D. R., *Interactive data analysis*. New York: Wiley, 1977.
- Merrill, M. D., *Adaptive models and learner control*. Working Paper No. 136. San Diego, Courseware Inc., 1979.
- Mosteller, F. & Tukey, J. W., *Data analysis and regression: A second course in statistics*. Reading, Mass.: Addison-Wesley, 1977.
- Opp, K. D., *Individualistische Sozialwissenschaft*. Stuttgart: Enke, 1979.
- Reither, F., *Der Einfluß der Selbstreflexion auf Strategie und Qualität des Problemlösens*. In H.-K. Garten (Hg.), *Diagnose von Lernprozessen*. Braunschweig: Westermann, 1977. S. 137—151.
- Reither, F., *Über die kognitive Organisation bei der Bewältigung von Krisensituationen*. In H. Ueckert & D. Rhenius (Hg.), *Komplexe menschliche Informationsverarbeitung*. Stuttgart: Huber, 1979. S. 210—222.
- Rey, W. J. J., *Robust statistical methods*. Berlin: Springer, 1978.
- Ritzen, J. M., Winkler, D. R. & Hargreaves-Heap, S., *Teacher preferences and the level and distribution of scholastic achievement*. *Journal of Experimental Education*, 1979, 47, 311—319.
- Scheuerman, K., *An inventory of adapting variables*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1976.
- Schwarzer, R., *Remedialer und adaptiver Unterricht*. *Unterrichtswissenschaft*, 1977, 4, 333—345.
- Seidman, E., Linney, J. A., Rappaport, J., Herzberger, S., Kramer, J. & Alden, L., *Assessment of classroom behavior: a multiattribute, multisource approach to instrument development and validation*. *Journal of Educational Psychology*, 1979, 71, 451—464.
- Simons, H. & Möbus, C., *Veränderung von Berufschancen durch Intelligenztraining*. Heidelberg: Bericht aus dem Psychologischen Institut der Universität, Diskussionspapier Nr. 8, Juli 1977.
- Snow, R. E., *Theory construction for research on teaching*. In M. W. Travers (Hg.), *Second Handbook of Research on Teaching*. Chicago, Rand McNally, 1973. S. 77—121.
- Snow, R. E., *Learning and individual differences*. In L. S. Shulman (Hg.), *Review of research in education*. Vol. 4. Itasca, Ill.: Peacock, 1977. S. 50—105 (a).
- Snow, E. W., *Individual differences and instructional theory*. Paper presented at the Annual Meeting of the American Educational Research Association, New York, 1977 (b).

- Treiber, B., Qualifizierung und Chancenausgleich in Schulklassen. Frankfurt: Lang, 1980.
- Treiber, B., Chancenausgleich in Schulklassen. DFG-Abschlußbericht. Heidelberg: Psychologisches Institut der Universität (in Vorbereitung).
- Treiber, B., Weinert, F.E. & Groeben, N., Bedingungen individuellen Unterrichtserfolgs. Zeitschrift für Pädagogik, 1976, 22, 153—179.
- Tukey, J. W., Exploratory data analysis. Reading, Mass.: Addison-Wesley, 1977.
- Wainer, H., Robust statistics: A survey and some prescriptions. Journal of Educational Statistics, 1976, 1, 285—312.
- Weinert, F.E., Remediales Lehren und Lernen. In K.-J. Klauer & A. Reinartz (Hg.), Sonderpädagogik in allgemeinen Schulen, Bd. 9. Handbuch der Sonderpädagogik. Berlin: Marhold, 1977.
- Weinert, F.E., Programmatische Anmerkungen zur Überwindung allzu optimistischer und pessimistischer Sackgassen in der Bildungsforschung. Bildung und Erziehung, 1980, 33, 39—43.
- Weinert, F.E. & Petermann, F., Erwartungswidrige Schülerleistungen oder unterschiedlich determinierte Schulleistungen? In H. Heckhausen (Hg.), Fähigkeit und Motivation in erwartungswidriger Schulleistung. Göttingen: Hogrefe, 1980.
- Weinert, F.E., Schneider, W. & Treiber, B., Bedingungsanalyse von (erwartungswidrigen) Schulleistungen. In H. Heckhausen (Hg.), Fähigkeit und Motivation in erwartungswidriger Schulleistung. Göttingen: Hogrefe, 1980.
- Westmeyer, H., Die rationale Rekonstruktion einiger Aspekte psychologischer Praxis. In H. Albert & K. H. Stapf (Hg.), Theorie und Erfahrung. Stuttgart: Klett-Cotta, 1979. S. 139—162.
- Wottawa, H., Systematische Agglutination von Hypothesen — Eine Ergänzung herkömmlicher Auswertungsverfahren. In MESA — Methoden und Modelle zur Erfassung des Streßverhaltens in Arbeitssystemen, Rpot 78.02. Heidelberg: Klinikum der Universität Heidelberg, 1978.
- Wottawa, H., Grundlagen und Probleme von Dimensionen in der Psychologie. Psychologia Universalis, 1979, 40, 1—122.
- Wottawa, H., Härtner, R. & Mattes, K., Computerunterstützte Hypothesenagglutination zur Erfassung komplexer Zusammenhänge. EDV in Biologie und Medizin, 1980.

Anschrift der Verfasser:

Dr. Bernhard Treiber, Dr. Wolfgang Schneider
Psychologisches Institut der Universität
Hauptstr. 47—51, 6900 Heidelberg