

BERICHT  
aus dem  
PSYCHOLOGISCHEN INSTITUT  
DER UNIVERSITÄT HEIDELBERG

Z U R INTERAKTION  
VON LERNMERKMALEN UND LEHRMETHODEN:  
REKONSTRUKTION UND-NOMIERUNG  
DES ATI-FORSCHUNGSPROGRAMMS

Bernhard Treiber u. Franz Petermann

Diskussionspapier Nr. 4

April 1976

Zur Interaktion  
von Lernermerkmalen und Lehrmethoden:  
Rekonstruktion und Normierung  
des ATI-Forschungsprogramms

Bernhard Treiber & Franz Petermann

Diskussionspapier Nr. 4  
April 1976

Eigentum des  
Psychologischen Instituts  
der Universität Gießen  
Hauptstadt 47-51

Infr I 17

ZUSAMMENFASSUNG

Die Problematisierung herkömmlichen Ansätze zur Erklärung und Vorhersage schulischen Lernens wird im ATI-Forschungsprogramm thematisiert, das die Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden auf den Lernerfolg analysiert und die Grundlage für schüleradaptive Lehr-Lern-Systeme bildet. Dieses Programm wird in seinem aktuellen Forschungsstand vorgestellt. Zur Erklärung seiner nur ungenügenden Leistungsfähigkeit wird sein bisheriger Verlauf rekonstruiert, wobei erhebliche Mängel im Heuristik-, Methodologie-, Erklärungs- und Anwendungsbereich erkennbar werden. Dem schließen sich Vorschläge zu seiner Optimierung auf unterschiedlichen Handlungsebenen an, die vor allem die Entstehung und weitere Entwicklung von ausdifferenzierten Lehr-Lern-Theorien ermöglichen sollen.

SUMMARY

Criticisms of traditional attempts at explaining and predicting school learning and scholastic achievement lead to the systematic analysis of attribute-treatment interactions which is also the basis for student-adapted teaching-learning systems. The current situation in this research program is described and found deficient. Thus, a reconstruction of its development up to now is carried out, revealing gross shortcomings in its heuristic, methodological, explanatory and technological dimensions. Suggestions are then made for optimizing future progress in these areas, with the main emphasis on initiating and stimulating the generation of elaborated teaching-learning-theories.

FUSSNOTE ZUM TITEL:

N. Groeben, W. Schneider & F.E. Weinert danken wir für kritisch-konstruktive Anmerkungen zur Erstfassung des Manuskriptes. A. Springer hat freundlicherweise die Darstellung eines mehrteiligen Flußdiagramms (Abb. 2) übernommen.

1. REKONSTRUKTION DER VORLÄUFER DES ATI-PROGRAMMS:  
VORHERSAGE UND ERKLÄRUNG VON SCHULLEISTUNGEN

Vorhersage und Erklärung individueller Schulleistungen erfolgten bislang in zwei unterschiedlich operierenden und organisierten Forschungsprogrammen (korrelativ vs. nomothetisch), die von verschiedenen Kernannahmen (z.B. über die relative Bedeutung von Schülermerkmalen vs. Lernumwelt- und prozeßbedingungen) ausgingen, dabei unterschiedliche (experimentelle vs. naturalistische) Prüfmodelle bevorzugten und sich mit ihren Ergebnissen für unterschiedliche Anwendungszwecke (z.B. Schülerselektion vs. Schul-, Unterrichtsplanung und -organisation) in variablem Ausmaß eigneten. (CRONBACH 1957).

Versuche, die individuelle Schulleistung durch Regression auf eine möglichst einfache und allgemein anwendbare Linearkombination von Schülermerkmalen hoher zeitlicher Stabilität vorherzusagen, gingen traditionell von dem Konzept der prädiktiven Validität (vgl. CRONBACH 1971) aus. In ihm wird ein interindividuell konstanter Einfluß von anderen bereits vorfindbaren Schülermerkmalen bzw. von den zwischen Prädiktor- und Kriteriumserhebung anzusetzenden Bedingungen des schulischen Lehr-Lern-Prozesses unterstellt, was auch die Ermittlung einer einheitlichen populations- und situationsinvarianten Regressionsgleichung rechtfertigt. Die Falsifikation dieser Konstanzannahmen erfolgte inzwischen durch den Nachweis, daß bei unterschiedlichen Schülergruppen und Lernumwelten (in Koeffizient und Konstante) sich unterscheidende Regressionsgleichungen zur Vorhersage der Schulleistungen resultieren (vgl. dazu SIMONS & MOBUS 1976), und initiierte Forschungsbemühungen, die differentielle Prognosevalidität der ursprünglich verwendeten Schulleistungsprädiktoren durch Erweiterung der Prädiktormenge vor allem um differentialpsychologisch bedeutsame Schülermerkmale zu reduzieren (zur Diskussion über Moderatorvariablen vgl. ABRAHAM & ALF 1972 ab, AMELANG & VAGT 1970, BÄRTUSSEK 1970, American Educational Research Journal Debate 1974, 11, 63-75, DUNNETTÉ 1972, GHISELLI 1972, JÄGER 1974, LISSITZ & SCHOENFELDT 1974, VELICER 1972, ZEDECK 1971). Diese Versuche, mit zunehmend mehr Schülermerkmalen eine zunehmend genauere und populationsunspezifische Vorhersagbarkeit von Schulleistungen zu erreichen (vgl. GAEDICKE 1974), weisen aber - z.B. schon wegen der Varianzüberlappung ihrer Prädiktoren und verschiedener Unzulänglichkeiten der entsprechenden Erhebungsinstrumente (etwa bei nicht-kognitiven Variablen) - einen

symptotisch abnehmenden Optimierungszuwachs auf, der auch den praktischen Nutzen dieser Forschungsstrategie für schulinterne Selektionszwecke in Frage stellt. Neuere Forschungsarbeiten analysieren denn auch zunehmend häufiger den differentiellen Einfluß von Merkmalen der (schulischen) Lernumwelt auf das Zustandekommen von Schulleistungsunterschieden (vgl. SIMONS, WEINERT & AHRENS 1975), und indizieren insoweit eine Abkehr von einer schülerzentrierten Analyseperspektive bei Schulleistungsprognosen. *vgl. auch Weinert: Qualität d. Sch. 1975*

Ansätze, die individuelle Schulleistungsvarianz als Ergebnis eines institutionalisierten Lehr-Lern-Prozesses durch Merkmale der schulischen Lernumwelt (Schul- und Unterrichtsorganisation, Lehrmethoden, Lehrerpersönlichkeit etc.) aufzuklären, führten einmal zu (makroanalytischen) Effizienzvergleichen von unterrichtlich variierbaren Methoden-, Programm-, Medien- oder Personvariablen. Sie verwendeten dabei, um den Bewährungsgrad und Informationsgehalt entsprechender Erklärungshypothesen möglichst groß zu halten, in der Regel nur globale Mittelwertvergleiche zwischen einzelnen Bedingungsstufen ohne Berücksichtigung von Kontext- oder Schülermerkmalen. Auch das Forschungsprogramm zur Erklärung der Schulleistung durch systematische Variation von Segmenten der schulischen Lernumwelt ging also zunächst von der Annahme einer interindividuell konstanten Wirkung der manipulierten Umweltbedingungen unabhängig von anderen Populations-, Kontext- und Situationsparametern von Verlauf und Ergebnis schulischer Lernprozesse aus. Die Adäquanz dieser Konstanzannahme wird durch die Anhäufung insignifikanter oder doch praktisch irrelevanter Ergebnisse nur globaler Mittelwertvergleiche zumindest in Frage gestellt. Dies begünstigt Erklärungsversuche, Veränderungen in Lernkriterien durch systematische Variation zunehmend enger umschriebener und experimentell leicht manipulierbarere Merkmale von Instruktionsformen bei strikter Kontrolle möglicher Störfaktoren zu erreichen, zu der auch individuelle Schülerdifferenzen zählen. Die Beibehaltung von möglichst einfach strukturierten und erklärungskräftigen Annahmen über die Verursachung von Schulleistungsdifferenzen durch schulische Lehr-Lern-Bedingungen war indes nur bei Anwendung auf idealisierte Experimentalsituationen zu erreichen und führte, da Generalisationsversuche auf unterrichtsähnliche(re) Lehr-Lern-Situationen unterblieben oder scheiterten (vgl. SNOW 1974), zu reduktiven und anwendungsfernen (Teilbereichs- oder Miniatur)Theorien des schulischen Lernens (z.B. in Form von Imitations- oder Verstärkungsmodellen). Auf diese Konsequenz wurde dann mit der Erweiterung der Wenn-Komponente von explanativen Aussagen vor allem um weitere extern variierbare Lernbedingungen reagiert, was aber (bei gleichbleibender Dann-Komponente) deren Informationsgehalt einschränkte und gleichwohl die Fiktion ihrer Populationsabhän-

gigkeit aufrechterhielt. Deshalb untersuchen neuere Forschungsarbeiten zur Erklärung von Schulleistungsdifferenzen, ob und in welchem Ausmaß externe Lernbedingungen Verlauf und Ergebnis von schulischen Lernprozessen auf eine für unterscheidbare Schüler(gruppen) variable Weise beeinflussen (s. WEINERT 1965).

Herkömmliche Versuche, Schulleistungen durch Beschreibung der internen Merkmale eines Lernenden (z.B. anhand von intellektuellen Fähigkeiten, kognitiven Stilen, Entwicklungsbedingungen, Vorkenntnissen, Motiven, Interessen und Einstellungen) vorherzusagen bzw. durch Manipulation der externen Lernbedingungen (z.B. Lehrmethoden, Lernhilfen, -materialien, -zeiten) zu erklären, führten in ihren Ergebnissen zu einem nur suboptimalen Erkenntnisstand und befinden sich damit in einem Zustand, in dem eine reale Steigerung des Prognose/Erklärungserfolgs entsprechender Zusammenhangsbehauptungen zunächst durch interaktive Verknüpfung von Elementen der Variablenklassen "Schülermerkmal" und "Unterrichtsmethode" in deren Wenn-Komponente zu erwarten ist. Vorhersagen von Schulleistungen werden dabei die vermittelnde Funktion intervenierender Lehr-Lern-Prozesse, Erklärungsversuche die den Einfluß externer Lernbedingungen moderierende Bedeutung interner Lernvoraussetzungen bei Schülern zu berücksichtigen haben.

Der stagnierende Erkenntnisfortschritt beider Forschungsansätze, die sich anschließende wissenschaftsinterne Problematisierung einzelner ihrer Kernannahmen und Vorgehensweisen und die Explikation alternativer Lösungshorizonte haben aber auch die Legitimität der Anwendungszwecke und die Effizienz der Verwertungsformen, die Prognosen und Erklärungen von Schulleistungen im gegenwärtigen Bildungssystem erfüllen bzw. begründen, in Frage gestellt, die Perzeption bildungspolitischer Handlungsprobleme verändert und zur Ausarbeitung eines alternativen Normenangebotes geführt, aus dem sich konsequent auch andere Ziele und Lösungen ergaben.

Schulleistungsprognosen sollen Entscheidungen über den Zugang oder die Zuweisung von Schülern zu alternativen Lernwegen und Ausbildungsgängen (von der Sonder- bis zur Hochschule) regeln oder doch erleichtern. Da Vorhersagen des künftigen Lern- und Ausbildungserfolges aber nicht unabhängig von der Variabilität des zwischenzeitlichen Einflusses der schulischen Lernumwelt sind, kann eine als gerecht geltende und neutrale Anwendung des formalen Gleichheitsprinzips bei der Vergabe von Bildungs- (und damit von Berufs- und Lebens-)chancen nicht mehr auf der Grundlage lernumweltunabhängiger allgemeiner Regressionsgleichungen erfolgen (vgl. CLEARY et al. 1975).

Die Erklärung der Schulleistung durch Merkmale der (schulischen) Lernumwelt soll u.a. eine optimale Planung und Organisation institutionalisierter Lehr-Lern-Prozesse ermöglichen. Da die Bestimmung der Effektivität einzelner (z.B. didaktischer oder organisatorischer) Maßnahmen aber zu je nach Schülergruppe variablen Ergebnissen führt, bedeutet eine populationsunspezifische Optimierung von Unterricht, die über die Bereitstellung eines für alle Schüler gleichen Angebots an jeweils nur durchschnittlich besseren Lernmöglichkeiten erfolgt, daß einzelne Schüler (gruppen) zugunsten anderer benachteiligt werden.

Die herkömmlichen Formen der Selektion von Schülern und der Optimierung von Unterricht erweisen sich demnach als inadäquat. Folgerichtig initiierte und unterstützte die Kritik an der Gültigkeit ihrer zentralen Annahmen und Begründungen eine Anzahl von bildungspolitischen Reformansätzen, die zum Ziel haben,

- defensiv die bestehenden Selektionszwänge (zumindest in Teilbereichen des Schulsystems) zu lockern, einzelne Selektionsmechanismen zu liberalisieren und bestimmte Folgeprobleme (z.B. die Irreversibilität von Selektionsentscheidungen) abzuschwächen;
- offensiv die bisherige lernerfolgsrelevante Benachteiligung einzelner Schülergruppen durch eine flexiblere Organisation schulischer Lehr-Lern-Prozesse in Formen der inneren und äußeren Differenzierung zu reduzieren.

Diese Ansätze konvergieren also in einer lernumweltbezogenen Schülerplatzierung und schüleradaptierten Unterrichtsorganisation.

## 2. DAS ATI-PROGRAMM: ANNAHMEN, ADÄQUATHEITSBEDINGUNGEN UND AKTUELLER FORSCHUNGSSTAND

Die damit vorgegebene Problemstellung und der daraus ableitbare Bedarf an prognostischen und technologischen Problemlösungen für die Planung und Steuerung adaptiver Unterrichtssysteme und -prozesse initiierte nun die Ausarbeitung eines systematisierten Forschungsprogramms zur Entdeckung und Sicherung von Wechselwirkungen zwischen Schülermerkmalen und Unterrichtsmethoden (oder ATIS = Aptitude-Treatment-Interactions), in dem sowohl heuristische Modelle zur Generierung von Hypothesen über solche Wechselwirkungen, von formal-statistischen Verfahren zu ihrer Überprüfung wie auch von organisatorischen Formen zu ihrer technologischen Nutzung ausgearbeitet wurden. Die Analyse dieses Forschungsprogramms steht im Mittelpunkt dieser Arbeit.

Die allgemeine Zielvorgabe und Problemstellung des ATI-Programms besteht darin,

- solche quantitativen oder qualitativen Variablen ausfindig zu machen, die den Vorhersagewert von Prädiktoren oder den Erklärungswert unabhängiger Instruktionsbedingungen signifikant verändern;
- Art und Ausmaß zu bestimmen, in dem teilgruppenspezifische Regressionsparameter von der für eine Gesamtgruppe von Individuen erstellten Regressionsgleichung abweichen;
- den mediativen Einfluß dieser Variablen auf eine allgemeine Prädiktor-Kriterium-Beziehung durch gehaltvolle Theorien zu erklären;
- die Ergebnisse dieser Untersuchungen prognostisch für eine nicht-diskriminierende Vorhersage eines Kriteriums (Kriteriensatzes) anhand der Kenntnis von Prädiktor- und Moderatorvariablen anzuwenden;
- explanativ für die Erklärung der differentiellen Bedingungs-, Verlaufs- und Effektstruktur von (institutionalisierten) Lernprozessen zu verwenden;

technologisch für die Schaffung optimaler Lernbedingungen zu nutzen, unter denen Lernende den für die geeigneten alternativ organisierten Lernumwelten zugewiesen werden.

Zur Beurteilung der wissenschaftlichen wie praktischen Brauchbarkeit der dabei generierten Forschungsergebnisse können die (in Tab. 1 aufgeführten) Adäquatheitsbedingungen herangezogen und mit den bisherigen Ergebnissen relevanter ATI-Forschungsarbeiten verglichen werden, die auch von BERLINER & CAHEN (1973), BRACHT (1969, 1970), CRONBACH & SNOW (1969) und FLAMMER (1975a) zusammengefaßt wurden.

Hier Tab. 1 einfügen

Das ATI-Forschungsprogramm weist in diesem Soll-Ist-Vergleich eindeutig ungenügende Erkenntnisleistungen sowie ein deutlich stagnierendes Fortschrittmuster auf. Das daraus resultierende empirische Prognose-, Erklärungs- und Technologie-defizit hat zudem zur Folge, daß derzeit eine rationale Planung und Organisation schüleradaptierter Unterrichtsformen zur optimalen Erreichung gegebener Lehrziele auf der Grundlage der vorliegenden ATI-Befunde nicht erfolgen kann (vgl. KALLOS & LUNDGREN 1975).

### 3. REKONSTRUKTION DES ATI-PROGRAMMS

Von diesem Sachverhalt geht auch die folgende Analyse aus. Zu seiner Erklärung wird sie auf die Existenz von Forschungsproblemen hinweisen, die bisher zumeist nicht berücksichtigt oder nur ungenügend beschrieben, bearbeitet und geklärt wurden und die deshalb als Gründe für die defizitäre Leistungsfähigkeit des ATI-Programms anzuführen sind. Dazu wird im folgenden - entsprechend den forschungsinternen Bedingungen eines Erkenntnisfortschritts - zunächst die im ATI-Programm benutzte Heuristik rekonstruiert und kritisiert, und in einem zweiten Schritt wird dann auf Probleme seiner Methodologie bei der Planung und Herstellung geeigneter Prüfbedingungen

sowie zur Analyse und Bewertung von Prüfergebnissen eingegangen. Die Explikation dieser Probleme und ihrer Lösungsvoraussetzungen sollen dann eine mittel- bis längerfristige Optimierung dieses Forschungsprogramms ermöglichen.

#### 3.1 ATI-HEURISTIK

Die Generierung von ATI-Annahmen, die besagen, welche und wieviele Spalten-Zeilen-Kombinationen einer (prinzipiell unbegrenzt erweiterungsfähigen) Matrix von Lerner- und Lernumweltmerkmalen einen als entdeckenswert unterstellten differentiellen Verlauf und Erfolg von (schulischen) Lernprozessen erwarten lassen, ist die Hauptaufgabe der ATI-Heuristik. Die Bearbeitung dieser Aufgabe erfolgt nun, sofern a priori überhaupt explizite ATI-Hypothesen formuliert und bezüglich ihrer Herkunft markiert werden, auf der Basis intuitiv begründeter Evidenzargumente mit allenfalls lockerem und nur analogisierendem Bezug zum tradierten psychologischen Theorienbestand oder in explorativen a posteriori-Analysen von Untersuchungen, die anderen Problemstellungen galten und in deren Datenkollektionen nachträglich nach signifikanten Wechselwirkungen gesucht wird, so daß bei deren Entdeckung ein plausibler Interpretationskontext nachträglich erst hergestellt werden muß.

Zurecht wird die Anwendung einer nur intuitiv-induktiven und theorieleeren bis -armen Heuristik von mehreren Autoren (vgl. SCHWARZER & STEINHAGEN 1975:17, 19) denn auch als unergiebig bezeichnet. Dieser Einschätzung schließt sich in der Regel dann aber der Vorschlag einer stärker theoriegeleiteten ATI-Forschungsheuristik an, der folgendes zu unterstellen scheint:

- die Existenz eines ausformulierten, systematisierten und hinreichend gut bestätigten Konstrukt- und Theoriebestandes über die Bedingungs-, Verlaufs- und Effektstruktur von (schulischen) Lehr-Lern-Prozessen, auf die sich die abhängige Variable als Effekt-Indikator bezieht: (= Theorie über die AV-Ergebnisse?)

- die Existenz empirisch hinreichend verankerter Beschreibungsstrukturen für differentiell lernerfolgsrelevante Schülermerkmale (Moderator- oder Mediatorvariablen);
- die Existenz von trennscharfen Klassifikations- und Taxonomiekriterien für Unterrichtsmethoden, denen ein für verschiedene Schülergruppen differentieller Effektivitätswert zugeordnet werden kann.

Diese Kenntnisse werden also vorausgesetzt, und von ihrer besseren heuristisch-konstruktiven Nutzung verspricht man sich eine deutlich erleichterte Entdeckung und Sicherung bedeutsamer ATI-Ergebnisse. Hinweise jedoch, welche Wissensbestände damit angesprochen und wie diese sinnvollerweise auszuwerten und anzuwenden sind, unterbleiben in der Regel. Dies liegt nun vor allem daran, daß entsprechende Theoriegefüge derzeit gar nicht verfügbar sind, aus denen sich eine am Lernerfolg ablesbare Interaktion von internen Merkmalen des Lernenden und externen Lernbedingungen begründet vorhersagen ließe. Die bisherigen Ansätze zur Vorhersage oder Erklärung von Verlauf und Ergebnis schulischen Lernens wurden nämlich, wie bereits ausgeführt, in getrennten Forschungsprogrammen entwickelt, wählten dabei unterschiedliche Analyseebenen und -strategien und kamen zu Ergebnissen, die einer Integration in Methoden-Merkmal-Verknüpfungen nur schwer zugänglich sein dürften. Ansätze, allgemein- und differentialpsychologische Perspektiven der Vorhersage und Erklärung schulischen Lernens in einer Matrix von externen Lernbedingungen (Strukturparametern) und internen Lernermerkmalen (Individualparametern) zu integrieren, sind derzeit jedenfalls erst in programmatischen Zieldeklarationen zu erkennen (vgl. FISCHER 1974, SIXTL 1972, VALE & VALE 1969), denen ausformulierte, überprüfte und (teil)bestätigte Theorieaussagen noch nicht gefolgt sind (vgl. KALLOS 1975, WALBERG 1974, WALSH 1975).

Bei Absenz eines technologiegenerierenden und -begründenden

den nomologischen Theorienkontextes über die Binnenstruktur von Lernprozessen eröffnen sich der Konzipierung einer ATI-spezifischen Heuristik demnach zwei Möglichkeiten:

(a) Künftige Forschungsaktivitäten kommen verstärkt einer anwendungsfernen Entwicklung von (Teilbereichs)Theorien über Ausschnitte der Bedingungs-, Prozeß- und Effektstruktur schulischen Lernens zugute, die kurzfristig zumindest als kostenintensive Ressourcenbindung zu Lasten einer unmittelbar rentablen Technologieproduktion geht, längerfristig aber trotz des durch sie freigesetzten Erkenntnisüberschusses auch für praktische Anwendungsbedürfnisse als ertragreichste Forschungsinvestition betrachtet wird (s. VAN DER DAELE & KROHN 1975).

(b) Künftige Forschungsaktivitäten werden, da eine schul- und unterrichtsrelevante Entwicklung von ATI-Technologien als allenfalls partiell abhängig von der Entwicklung differenzierter Theorien des Lernens eingeschätzt wird, unter Verzicht auf einen weiterreichenden theoretischen Erklärungsanspruch begrenzt auf die Bestimmung nur des durch Schülermerkmale modifizierten Anwendungserfolges zweier oder mehrerer Unterrichtsmethoden zur Erreichung vorgegebener Lehrziele.

Das ATI-Program hat, begründbar vor allem mit dem erheblichen unterrichtspraktischen Informationsbedarf, der Komplexität ihres Gegenstandsbereichs und der fehlenden Problemlösungskapazität relevanter Theorieausschnitte, die zweite Strategie gewählt und trotz ungenügender Kenntnis der hypothetischen Parameterstruktur der zugrunde liegenden Lernprozesse zu bestimmen versucht, unter welchen Lernbedingungen (Unterrichtsvarianten) welche Schüler welchen Lernerfolg aufweisen. Dieser theoretische Versuch, unter einer vergleichsweise großen Anzahl möglicher Differenzierungsmerkmale im Lernumwelt- und Lernerbereich Variablenkombinationen von hinlänglicher Interaktionswahrscheinlichkeit ausfindig zu machen und in die Hypothesenformulierung einzubringen, erwies sich — angesichts der Viel-

zahl plausibler Merkmal-Methoden-Kombinationen - aber als durchweg ineffizient und unergiebig. Die - bezogen auf die eingangs (in Tab. 1) explizierten Adäquatheitskriterien - erfolglose Anwendung dieser Black-Box-Strategie problematisiert demnach die These, Theorie- und Technologieentwicklung in diesem Forschungsbereich seien nur lose miteinander verknüpft. Sie belegt vielmehr, daß eine weitere theorieabstinente Funktionalisierung des ATI-Programms zugunsten kurzfristiger Nutzungsansprüche dessen technologischen Innovationspotential zunehmend begrenzen und die systematisierte Suche nach und Entdeckung von bedeutsamen Merkmal-Methoden-Wechselwirkungen längerfristig erfolglos werden lassen würde.

Hinzu kommt, daß der Mangel an theoretischer Strukturierungsfähigkeit der extern vorgegebenen Problemstellung, die Planung, Organisation und Steuerung adaptiver Unterrichtsprozesse zu ermöglichen, durch deren Umformulierung in eine forschungspraktisch besser bearbeitbare aber zielabweichende Form zu kompensieren versucht wurde. In der Regel wurden nämlich zunächst mehrere (zumeist nur zwei) alternative Unterrichtsmethoden zur Erreichung eines Lernkriteriums aufgrund von Plausibilitätsmodellierungen ausgearbeitet, deren für unterscheidbare Schülergruppen möglichst diskrepanter Anwendungserfolg dann durch solche Merkmale vorherzusagen versucht wurde, die sich post hoc aus einer möglichst großen Anzahl simultan erfaßter Persönlichkeitsvariablen als hinreichend brauchbar erwiesen. Folglich wird die ATI-Forschungsperspektive, Instruktion an interne Bedingungen des Lernenden anzupassen, bei ihrer strategisch-operativen Konkretisierung auf eine Weise restringiert, in der nur mehr Schülermerkmale mit einem möglichst hohen Diskriminationswert für die Vorhersage des differentiellen Erfolges **b e r e i t s** ausgearbeiteter Unterrichtsmethoden (fixed treatments) selektiert werden. Das Forschungsproblem ist demnach nicht mehr, durch die Bereitstellung welcher und wie vieler Instruktionsvarianten ein ge-

gebenes Lernkriteriums durch möglichst viele Lernende erreicht werden kann, sondern nur noch, welche Schüler welchen bereits definierten Unterrichtsalternativen wie erfolgreich zugeordnet werden können. Durch die vorgängige Festlegung von Art und Anzahl der bereitgestellten Unterrichtsalternativen, die ohne explizite Berücksichtigung von Schülermerkmalen vor allem im Hinblick auf verfügbare Forschungsressourcen erfolgt, wird aber auch der intendierte Adaptivitätsnutzen zwangsläufig eingeschränkt.

Es ist schließlich zu vermuten, daß die Ergebnisse weiterer ATI-Untersuchungen bereits im Heuristikbereich an die Stelle des ursprünglich nur auf die Verknüpfung von Unterrichtsmethoden und Schülervariablen bezogenen ATI-Modells ein erweitertes Programm treten lassen werden, in dem die für die Vorhersage der differentiellen Lernleistung verschiedener Schülergruppen unterstellte Zusammenhangsstruktur von Merkmalen der Lernaufgaben (Tasks), des zu ihrer Lösung qualifizierenden Angebots an Unterrichtsalternativen (Treatments) und der ihre Anwendung und Nutzung steuernden Lehrpersonen (Teachers oder Traits) (=ATTI-Konzept) untersuchenswert erscheint (vgl. RHETTS 1974, WEBSTER & MENDRO 1974). Die Komplexität dieser Problemperspektive erreicht gegenwärtig noch keine unbekanntere Forschungsarbeit. Allenfalls ist die Inkonsistenz einfacher Interaktionseffekte erster Ordnung über verschiedene Prüfinstanzen hinweg als Indikator für das Vorliegen komplizierterer Interaktionsbeziehungen höherer Ordnung zu werten, von dem die Auftretenswahrscheinlichkeit "einfacher" ATI-Befunde abhängt. Darüberhinaus weisen nur vereinzelte Untersuchungen ausschnitthaft auf den aufgaben-(vgl. RHETTS 1974) und lehrerspezifischen (vgl. WEBSTER & MENDRO 1974) Geltungsbereich von "einfachen" ATI-Befunden hin und schränken damit den möglichen Generalisationsradius weitaus enger ein als dies in bisherigen Forschungsaktivitäten (vgl. CRONBACH 1975) ex-



plizit gemacht wurde. Die mit der Expansion und Elaboration einer ATI- zu einer ATTTI-Perspektive vollzogene Komplexitätssteigerung bereits in der Formulierung der Problemstellung und die Reziprok dazu abnehmende Strukturierungsfähigkeit der herkömmlichen Lehr-Lern-Forschung, in der die entsprechenden Theorieelemente, Erhebungs- und Analyseinstrumente, Prüfmodelle, Datenverarbeitungsmethoden und Interpretationsangebote noch nicht bereitgestellt werden können, erhöhen demnach zwangsläufig die kognitive Resistenz des ATI-Forschungsprogramms gegenüber dem externen Nutzungsbedarf an curricularen Planungs-, Implementations- und Steuerungsalten.

Die sich daraus ergebenden steigenden Belastungen für die Problemlösungsfähigkeit des ATI-Programms sind möglicherweise aber auch durch Fokussierung der Grundlagenforschung auf jene Theorielücken nicht abzubauen, deren Schließung als Voraussetzung für den Erklärungserfolg von Theorien schulischer oder schulnaher Lehr-Lern-Prozesse angesehen wird. Kann dieser Komplexitätszuwachs, wie er mit der (prinzipiell unbegrenzten) Ausweitung der ATI-Analyseperspektive über die Entdeckung und Sicherung nur einfacher Interaktionseffekte hinaus (vgl. CRONBACH 1975) verbunden ist, nicht mehr konstruktiv abgefangen werden, wäre die Adäquanz einer konsequenten Nomologisierungstrategie zumindest für Teilbereiche psychologischen Theoretisierens in Frage gestellt, deren Ziel, hinreichend bewährtes Prognose- und Erklärungswissen mit möglichst großem Geltungsbereich und hohem Informationsgehalt bereitzustellen, als illusionär aufzugeben.

Daraus können derzeit unterschiedlich weitreichende Konsequenzen gezogen werden; einige davon lassen sich wie folgt typisieren:

- Die Möglichkeit hochkomplexer Interaktionsbeziehungen wird zwar gesehen, (wegen des zu erwartenden nur geringen Vorhersage/Erklärungszuwachses oder des erforderli-

chen Forschungsaufwandes) aber als nicht weiter untersuchenswert oder überprüfbar eingeschätzt oder als eindeutig unerwünscht zu ignorieren versucht (Problemverdrängung).

So eröffnet eine Transformation von Rohdaten die Möglichkeit, das Auftreten unerwünschter Interaktionen zwischen unabhängigen Variablen zu unterbinden (vgl. HUCK & SUTTON 1975, WINER 1971), wenn individuelle Unterschiede als für die Formulierung, Überprüfung und Geltung nomothetischer Aussagen irrelevante Fehlervarianz eingestuft werden.

- Die Möglichkeit von Interaktionseffekten auch höherer Ordnung wird akzeptiert; es wird deshalb versucht, nur noch eng umschriebene Selegate des jeweiligen Objektbereichs durch ad hoc oder analogisierend konstruierte Miniaturtheorien unter kontrollierten Bedingungen zu beschreiben und zu erklären (konventionelle Problembearbeitung). Dahinter steht zuweilen die Annahme, erfolgreiche bewährte Teilbereichstheorien liessen sich nachträglich immer noch in umfassendere und allgemeinere Theoriesysteme transformieren, in denen sie als Vorgänger-Theorien logisch bereits enthalten sind und, in Abhängigkeit von der variablen Transitivität intertheoretischer Reduktionsrelationen, einen mehr oder minder weitreichenden kumulativen Erkenntniszuwachs ermöglichen.
- In metatheoretischen Rekonstruktionen wird der gleiche Sachverhalt hingegen nur als Indikator gesehen für ein strukturell bedingtes allgemeines Stagnationsmuster von Theorienbildung und Erkenntnisfortschritt einer nomothetisch konzipierten Psychologie, das ohne Kritik und Veränderung seiner theoretisch-methodologischen Determinanten nicht aufzubrechen ist. Argumente gegen eine nomologisch orientierte Theorienbildung und ein empirisch-analytisches Methodologieprogramm verweisen dabei auf die allmählich unübersehbare Parzellierung der mit ihrer Hilfe analysierten Problembereiche, und plädieren wegen der epistemolo-

gisch wie pragmatisch unbefriedigenden Konsequenzen für die Erprobung anderer (z.B. diagraphischer) Wissenschafts- und (z.B. hermeneutischer oder phänomenologischer) Methodologiekonzeptionen in der Psychologie (Problemveränderung).

Der Erfolg dieses zuletzt genannten Versuchs, die geringe Leistungsfähigkeit des ATI-Programms als unvermeidbare Konsequenz seiner theoretisch-methodologischen Restriktionen zu erklären, hängt deshalb entscheidend von den Ergebnissen der Rekonstruktion dieses Programms ab und der Möglichkeit, konstruktive Hinweise für die Optimierung seiner künftigen Leistungsfähigkeit ausfindig zu machen und zu nutzen.

### 3.2 ATI-METHODOLOGIE

Kernannahme des bisherigen ATI-Forschungsprogramms ist die differentielle Vorhersagbarkeit von Kriteriumsleistungen durch Lernermerkmale bei unterschiedlichen Lehr- oder Unterrichtsformen. Es ist demnach von folgenden analytischen Elementen auszugehen: *Operationalisierung von:*

- (1) von einem Kriterium oder Kriteriensatz, das (der) hinreichend objektiv, reliabel und (lehrziel)valide erfaßt werden kann: *(1) cut score*
- (2) von  $n \geq 2$  alternativen Unterrichtsmethoden, deren Anzahl festgelegt ist und deren kontrollierte Realisierung gesichert werden kann: *(2) treatment*
- (3) von  $n \geq 1$  Schülermerkmalen, die einen möglichst hohen Diskriminanzwert für die Zuordnung von (bereits vorgegebenen oder erst noch zu konstruierenden) Unterrichtsalternativen zu unterscheidbaren Schülergruppen oder -quoten aufweisen. *(3) aptitude*

#### 3.21 Explikation und Operationalisierung der ATI-Elemente

Es wird im folgenden zu analysieren sein, ob und in welchem Ausmaß diese Voraussetzungen gegeben sind und zu welchen Konsequenzen Einschränkungen hinsichtlich ihrer Explikation und Operationalisierung führen können (vgl. dazu BRACHT 1969, 1970).

*int. Val. ?  
ext. Val. ?*

#### 3.211 Kriteriumserfassung

BRACHTs (1969) Übersicht ist zu entnehmen, daß in 100 von 108 der von ihm reanalysierten Fälle eine instruktions-spezifische und einfach strukturierte Lernleistungsvariable das Kriterium darstellte, oft auch laborexperimentell standardisierte Indikatoren der individuellen Leistungsfähigkeit in typischen (z.B. paarassoziativen, seriellen, frei reproduzierenden) Lern- und Gedächtnisaufgaben. Die Präferenzierung dieser voraussetzungsarmen-Kriteriumsvariablen von zumeist nur geringer Referenzgeneralität (vgl. SNOW 1974) und ihre Erhebung unter relativ kurzen Veränderungsintervallen sind mit der Erhöhung der internen Validität solcher Prüfversuche zu begründen, die gleichwohl mit einer abnehmenden Relevanz für schulische Anwendungsbereiche und Lernaufgaben bezahlt wird. Über die Erfüllung der üblichen Gütekriterien in diesen zumeist nur ad hoc konstruierten Erhebungsinstrumenten ist jedoch kaum etwas bekannt; desgleichen fehlt die Angabe von Beobachtungstheorien für die Validierung von Inferenzschlüssen über die Indikator-Konstrukt-Beziehung, sofern mit der abhängigen Variablen überhaupt eine latente Fähigkeitsdimension oder ein hypothetischer Lernprozess intendiert wird. Existenz und Definition der zu messenden Dimension werden dann aber bereits vorausgesetzt und also nicht von einer vorgängigen Konstrukt- und Theorienbildung über den intendierten Objektbereich abhängig gemacht (vgl. FISCHER 1974). Ebenso wird eine multikriteriale Verankerung eines Fähigkeits- oder Leistungskonstrukts durch Zuordnung von mehreren Indikatorvariablen vermieden; selbst in Fällen, in denen mehrere abhängige Variablen erhoben wurden (vgl. z.B. LOTT & BERNICE 1966, MERRILL & STOLUROW 1966), wurden ATI-Annahmen jeweils getrennt überprüft, blieb die Möglichkeit ihrer multivariaten Testung ungenutzt.

Annahmen über Art und Verlauf der durch unterschiedliche Instruktionsmaßnahmen angestrebten Veränderung von Schüler-Fähigkeits-Dispositionen sind aber bereits hinsichtlich der Wahl des optimalen Erhebungszeitpunktes für die abhängige Variable erforderlich (vgl. GALTUNG 1967, RENN 1973). Wird diese unmittelbar nach einem vergleichsweise kurzen Instruktionsintervall erhoben, bildet sie möglicherweise den Methodeneffekt noch nicht oder eben nur unzureichend ab. Bei einem längeren Instruktionsintervall werden die Meßergebnisse aber möglicherweise durch verschiedene Fremdeinflüsse verzerrt, so daß sie für eine reliable Schätzung des Methodeneffektes gleichfalls ungeeignet sind. Hinzu kommt, daß der differentielle Einfluß von Instruktionsvarianten und Lernermerkmalen auch zu unterschiedlichen Verlaufsmustern eines Lernprozesses führen kann (vgl. FLAMMER 1975\*) und sich deshalb auch zu verschiedenen Zeitpunkten unterschiedlich stark bemerkbar macht. Null-differenzen zum Zeitpunkt  $t_1$  sind möglicherweise überzufällig bedeutsame Unterschiede in  $t_{i+1}$ .

Die Adäquanz der Kriteriumserfassung ist aber auch durch folgenden Umstand eingeschränkt: die Erhebung der abhängigen Variablen soll in ATI-Untersuchungen eine brauchbare Schätzung eines (differentiellen) Lehreffektes ermöglichen, der in der Veränderung einer individuellen Fähigkeit seit Instruktionsbeginn besteht, eine bestimmte Aufgabenmenge erfolgreich zu lösen. Die Messung einer solchen Veränderung wirft nun nur solange kein Problem auf, wie eine randomisierte Zuteilung von Vpn auf die einzelnen Instruktionsvarianten des Methodenfaktors möglich ist: nur dann nämlich erlauben die Posttestwerte eine von systematischen Fehlern (und d.h. vor allem von der Schüler-Start-Fähigkeit zu Beginn einer Unterrichtseinheit) (weitgehend) unverzerrte und damit intern valide Schätzung des Instruktionseffektes. Ist aber eine solche zufällige Aufteilung (vor allem unter schulischen oder schulähn-

lichen Untersuchungsbedingungen) nicht durchführbar, werden Lernfortschritt und Lehreffekt auch in ATI-Untersuchungen durch Veränderungsmaße zu erfassen versucht, die den Einfluß von Prätest-Differenzen zu Beginn einer Unterrichtseinheit auf die Posttest-Werte nachträglich zu berücksichtigen versuchen, die wichtigsten Probleme einer brauchbaren Veränderungsmessung aber nur ungenügend lösen. Dies sind im einzelnen (vgl. auch Tab. 2):

Hier Tab. 2 einfügen

- (a) Prozentzuwachswerte (Gewinnprozente) sind in ihren verschiedenen Formeln durchweg ungeeignet als Veränderungsmaß, da ihre Anwendung nicht nur eine Intervallskala, sondern auch einen absoluten Nullpunkt voraussetzt (vgl. KLAUER 1969: 54). Die Abhängigkeit der Gewinnprozente von ihren Ausgangswerten führt im übrigen zu Regressionseffekten.
- (b) Differenzwerte ergeben sich aus der Differenz zwischen Prä- und Posttestwert (vgl. O'CONNOR 1972). Dabei wird jedoch übersehen, daß Differenzwerte, da in sie die Meßfehler ihrer Ausgangswerte eingehen, besonders unrelibel sind und ferner von ihren Prätestwerten abhängen und deshalb zu Regressionseffekten führen (vgl. CRONBACH & FURBY 1970, LORD 1963).
- (c) Kovarianzanalytisch korrigierte Posttestwerte ergeben sich aus der Projektion der bivariaten Verteilung der Vorher-Nachher-Werte auf die Koordinate der Posttestwerte (vgl. FAHRENBERG 1967). Hierbei wirken sich jedoch die Anfangsunterschiede zwischen den verschiedenen Gruppen mit ihrer vollen Artefaktträchtigkeit aus. Eine (in ATI-Arbeiten ungenutzte) Weiterentwicklung und Verbesserung beschreiben jedoch schon LORD (1958, 1963) und CRONBACH & FURBY (1970) sowie RENN (1974:66).
- (d) Residualwerte ergeben sich aus der Differenz zwischen tatsächlichem und dem (anhand des Prätestwertes und aus der Kenntnis der Prä-Posttest-Regression) geschätzten Posttestwert, werden aber vor allem wegen ihrer mangelnden Reliabilität kritisiert (vgl. O'CONNOR 1972).

Die ungenügende Korrektur des Einflusses von Prätest-Differenzen auf die Posttestwerte sowie die mangelnde Reliabili-

tät der genannten Veränderungsmaße beeinträchtigen folglich auch eine brauchbare Schätzung von Lehreffekten, ohne daß das Ausmaß der dadurch bedingten Verzerrungen im einzelnen zu beurteilen wäre. Sie bedarf deshalb der Kontrolle. Dazu bietet sich zum einen die in zahlreichen ATI-Studien vorgenommene randomisierte Aufteilung von Vpn und Erhebung nur von Posttestwerten an, die allerdings pragmatischen Restriktionen schulischer bis schulnaher Untersuchungssituationen unterliegt. Kann diese Voraussetzung nicht erfüllt werden, eröffnen lehreffektorientierte Tests mit Änderungssensitiven Tests sowie probabilistisch konzipierte lehrzielorientierte Tests eine wenn auch aufwendige konstruktive Möglichkeit einer verbesserten Veränderungsmessung (vgl. PETERMANN & KNOPF 1976). Vor allem die zuletzt genannte Möglichkeit einer probabilistischen Messung der abhängigen Lernvariable (z.B. nach Modellen der RASCH-Theorie) erfordert aber (vgl. FISCHER 1974) die Verfügbarkeit eines hinreichend formalisierten Lernmodells,

- dessen Kernannahmen (z.<sup>B.</sup> über Geltungsbereich und Verlaufsmuster) psychologisch begründet sind,
- das empirische Geltung für eine gegebene Aufgabenmenge und Schülergruppe aufweist,
- und das den Einfluß individueller Unterschiede (z.B. individuell unterschiedlicher Lehreffekte im ATI-Sinne) auf die allgemeine Modellstruktur thematisiert und präzisiert, d.h. entweder ablehnt oder durch Individualparameter berücksichtigt.

Liegt ein solches Lernmodell vor, so ist eine psychologisch verankerte und methodisch befriedigende Darstellung der aus unterschiedlichen instruktionsmaßnahmen resultierenden Änderungen des Ausprägungsgrades einer oder mehrerer latenter Schülerdispositionen möglich, die in herkömmlichen Messungen im Rahmen der klassischen Testtheorie nicht gelingt.

### 3.212 Prädiktorvariablen

Auswahl, Explikation und Erhebung, der in ATI-Untersuchungen verwendeten Schüler- oder Lernermerkmale unterstellen, daß durch diese Indikatoren wichtige Bedingungen differentieller Prozeß- und Effektstrukturen instruktionsbedingter Veränderungen einer latenten Schülerfähigkeit hinreichend gut erfaßt werden können. Bei fehlender oder ungenügender Kenntnis auch nur des allgemeinen Bedingungsgefüges typisierbarer Lehr-Lern-Situationen, von der aber gegenwärtig in diesem Forschungsbereich auszugehen ist, bleibt die Erstellung der Prädiktionsgleichung auf Plausibilitätsüberlegungen verwiesen, an denen sich die Nicht-Berücksichtigung wie Verwendung einzelner Lernermerkmale als Prädiktoren orientiert. Das Risiko, dabei bedeutsame Differentialmerkmale zu übersehen oder unwichtige Variablen als untersuchenswerte Prädiktoren aufzunehmen (Spezifikationsfehler) (vgl. BOHRNSTEDT & CARTER 1971), läßt sich bei einem nur alltagstheoretisch fundierten Vorgehen nicht reduzieren; allenfalls kann durch Erhebung einer möglichst großen Anzahl von Variablen die Wahrscheinlichkeit, wichtige Variablen zu übersehen, gesenkt werden, was besonders in explorativen Entwicklungsphasen von Bedeutung ist.

Die Entdeckungschance von Prädiktoren mit möglichst hohem Diskriminanzwert wird aber noch durch folgenden Umstand eingeschränkt: Die bisher entwickelten differentialpsychologischen Beschreibungs- und Erklärungsstrukture, deren Validität durch das Vorliegen eines hinreichend engmaschigen nomologischen Netzwerkes bereits als gesichert gelten darf, umfassen in der Regel eine größere Anzahl von theoretisch wie empirisch verknüpften Sachverhalten, denen sie eine (mehr oder minder) gemeinsame konstruktsspezifische Bedeutung verleihen (z.B. im Sinne der "Intelligenz", "Leistungsmotivation", "Ängstlichkeit" etc.), die sich zudem als hinreichend stabil gegenüber Zeit-, Populations- und Situationsunterschieden sowie gegenüber einer Substitution ihrer Indikatorvariablen erweisen (vgl. HERRMANN 1969). Insofern zielen diese Konstrukte

auf ein möglichst hohes Maß an Generalität, Universalität und Stabilität. Diese Merkmale schränken aber zugleich auch die Brauchbarkeit der genannten Konstrukte im Rahmen der ATI-Forschung ein: jener Grad an Spezifität, wie er für die Vorhersage unterschiedlicher Verlaufsmuster von Lehr-Lern-Prozessen und differentieller Effektivitätswerte alternativ organisierter Instruktionsprozesse erforderlich ist, wird von ihnen nicht erreicht (vgl. bereits CRONBACH & GLESER 1957: 125-127, SALOMON 1973). Von 74 aus insgesamt 108 ATI-Kombinationen der von BRACHT (1969) zusammengestellten 90 Einzeluntersuchungen, die als Beschreibungsindikatoren für faktoriell komplexe Konstrukte aufzufassen sind, wird denn auch nur in einem Falle eine signifikante disordinale Interaktion berichtet.

Werden andererseits instruktions-, lernprozeß- und lehrzielspezifische Prädiktoren gewählt, die einen höheren Diskriminanzwert als globale Persönlichkeitsmerkmale versprechen, so bedürfen Explikation und Operationalisierung dieser ad hoc konstruierten Beschreibungsindikatoren sowie ihre Zuordnung zu bereichsspezifischeren Konstrukten erst noch der Bestätigung durch geeignete Beobachtungstheorien, die jedoch entweder nicht vorliegen, nicht aufgeführt werden oder bislang ohne Bewährung blieben. Gütekriterien für diese Erhebungsinstrumente werden jedenfalls kaum genannt.

### 3.213 Unterrichtsmethoden

Probleme der Konzipierung und Realisierung alternativer Lernbedingungen sind weitgehend ATI-unspezifisch; sie treten auch bei einfachen Effektivitätsvergleichen von Unterrichtsmethoden, Material-, Medien- oder Programmvariablen auf und brauchen deshalb nur noch summarisch erwähnt zu werden.

Methodenstudien setzen üblicherweise voraus,

- daß vergleichbare Effektivitätskriterien vorliegen, die einer Messung zugänglich sind;
- daß Effektivitätsstudien unter vergleichbaren Bedingungen

durchgeführt werden;

- daß eine hinreichend eindeutige Beziehung zwischen der Variation von Methodenmerkmalen und dem Lernerfolg nachweisbar ist.

Diese Minimalvoraussetzungen sind aber nur selten gegeben (vgl. z.B. ROTH 1971: 20ff.).

Zum einen ist nämlich die Anwendung unterschiedlicher Methoden auch mit mehr oder minder variierenden Zielvorstellungen verknüpft, die eine Kriteriumsspezifische Explikation erfordern. Zudem muß ohne Analyse der tatsächlichen Implementation einer Unterrichtsmethode angenommen werden, daß sich trotz einer homogenisierbaren Zielkonzeption innerhalb einzelner Lehrmethoden ihre unterrichtliche Realisierung etwa durch verschiedene Lehrer unterscheidet, was die Entdeckung von Lehreffekten in Abhängigkeit von der Größe dieser Binnenvarianz erschwert bis verhindert.

CRONBACH & WEBB (1975) reanalysierten eine frühe ATI-Untersuchung, in der eine signifikante Interaktion berichtet wurde zwischen der Schüler-Unterscheidung als "Under- vs. Overachiever" und dem Ausmaß, in dem im Instruktionsprozeß der Bedeutungsgehalt des Lehrstoffs akzentuiert wurde. Sie fanden, daß diese Signifikanz bei Berücksichtigung der Klassenzugehörigkeit der einzelnen Schüler, die in der Originalarbeit unterlassen wurde, nicht mehr zu bestätigen war; allerdings mußte auf einen Vergleich der klassenspezifischen Regressionen wegen verschiedener Design- und Datenmängel verzichtet werden. Die Autoren geben indes zu bedenken, daß der Instruktionsprozeß bei unterschiedlichen Lehrern trotz deren Zugehörigkeit zu einer bestimmten Bedingungsstufe des Methodenfaktors mehr oder minder variiert (Binnenvarianz) und damit auch die Entdeckungschance eines Interaktionseffektes mehr oder minder erschweren kann.

Schließlich erweist sich eine eindeutige Aufklärung lehrmethodenspezifischer Effekte bei der Vielzahl simultan einwirkender Einflußfaktoren umso schwieriger, je weniger diese bekannt und deshalb kontrollierbar sind.

Die häufig berichtete Insignifikanz und praktische Irrelevanz einer Variation von Lehrmethoden geht im wesentlichen auf diese Probleme zurück und verweist mithin auf einen erheblichen Erklärungsabstand, der zwischen den unabhängigen Methoden-, Programm- oder Materialvariablen und davon abhän-

gigen Lernvariablen besteht und der ohne Berücksichtigung des zugrunde liegenden Lernprozesses, der Sachstruktur der zu bearbeitenden Aufgabenmenge sowie der im Instruktionsprozeß zu verändernden hypothetischen Schülerfähigkeit nicht reduziert zu werden vermag. Dies gilt auch für die Explikations- und Realisationsmöglichkeiten des Methodenfaktors in ATI-Untersuchungen (vgl. HEYMAN 1975).

### 3.22 Überprüfung von ATI-Aussagen

Die formale Bereitstellung der Elemente von ATI-Aussagen bildet nur eine notwendige Voraussetzung für die inhaltlich bestimmte Zuordnung geeigneter Prädiktorvariablen und Unterrichtsvarianten zur differentiellen Vorhersage gegebener Kriterien sowie die empirische Überprüfung und Bestätigung ihres interaktiven Zusammenhanges. Die nur unzureichende Erfüllung der genannten Kriterien erklärt denn auch noch nicht vollständig die unbefriedigenden Ergebnisse bei der Überprüfung von ATI-Hypothesen. Im folgenden wird deshalb eine Rekonstruktion der Prüfbedingungen für ATI-Annahmen unternommen, die in der Regel in zwei Schritten erfolgt:

- (1) durch den Nachweis eines signifikanten Interaktionseffektes gegebener Größe,
- (2) durch die Klärung, ob ein gegebener Interaktionseffekt als disordinal (, semi-disordinal) oder ordinal zu bezeichnen ist.

Beibehalten werden kann eine ATI-Annahme dabei erst bei Disordinalität des Interaktionseffektes. (Das Problem, ob und welche Bedeutung nur signifikant ordinale Wechselwirkungen haben können, wird weiter unten diskutiert) (vgl. 3.4).

Die dafür erforderlichen (vor allem statistischen) Prüfprozeduren wurden bereits ausführlich von anderen Autoren dargestellt, so daß darauf nicht näher eingegangen zu werden braucht (vgl. BRACHT 1969, 1970; BREDEKAMP 1975, PLOMP 1974).

Die Überprüfung von ATI-Annahmen erfolgte bislang zumeist in varianzanalytischen Untersuchungs- und Auswertungsplänen durch Kombination einer (in der Regel auf zwei Bedingungsstufen begrenzten Variation der ) Unterrichtsmethode mit einem oder mehreren (gleichfalls zumeist nur zwei- bis drei-

stufigen) Schülermerkmal(en). Die Identifikation einer Merkmals-Methoden-Interaktion hängt deshalb von der Signifikanz des F-Wertes für den Interaktionsausdruck ab; die Ablehnung der statistischen Nullhypothese und die Annahme der Alternativhypothese begründen die Beibehaltung einer ATI-Hypothese (vgl. LEE 1961)

Das Risiko einer Fehlentscheidung wird dabei durch  $\alpha$  (fälschliche Ablehnung der  $H_0$ ) und  $\beta$  (fälschliche Ablehnung der  $H_1$ ) beschrieben. Die Güte einer Untersuchung kann deshalb auch innerhalb des ATI-Programms danach beurteilt werden, ob die  $\alpha$ - und  $\beta$ -Wahrscheinlichkeit bekannt sind und vergleichbar klein gehalten werden können. Erst dann nämlich ermöglicht die Ablehnung einer der beiden statistischen Hypothesen ( $H_0$ ,  $H_1$ ) auch die begründete Annahme ihrer jeweiligen Alternative bei jeweils bekanntem Fehlerrisiko.

Auch in ATI-Untersuchungen wird aber nun, wie auch in anderen Forschungsarbeiten,  $\alpha$  bereits im voraus explizit festgelegt, während  $\beta$  unbekannt ist. Dennoch kann über Reanalysen auch nachträglich noch versucht werden, die  $\beta$ -Fehlerwahrscheinlichkeit bei der Ablehnung von ATI-Annahmen über Interaktionseffekte variabler Größe zu rekonstruieren.

So zeigt TREIBER (1976) für 77 der von BRACHT (1969) zusammengestellten ATI-Untersuchungen, daß die Wahrscheinlichkeit einer fälschlichen Ablehnung einer ATI-Annahme zugunsten der sie falsifizierenden Nullhypothese im Durchschnitt nur dann hinreichend klein ( $\beta \leq .20$ ) gehalten wurde, wenn zugleich anzunehmen war, daß die zu entdeckenden Methoden-Merkmal-Wechselwirkungen einen vergleichsweise großen Anteil ( $w^2 \geq .20$ ) an der Gesamtvarianz abhängiger Variablen aufklären. Nur in diesem Effektgrößenbereich konnte deshalb in den reanalytierten Arbeiten eine für ATI-Annahmen falsifizierende Konsequenz gezogen werden. Für kleinere Effekte ( $w^2 \leq .20$ ) mußte hingegen für die gleiche Folgerung eine weitaus höhere Fehlerwahrscheinlichkeit in Kauf genommen werden, so daß es unzulässig erscheint, aus der Nicht-Ablehnung der Nullhypothese auch auf deren Zutreffen zu schließen. Als Falsifikationsinstanz für ATI-Annahmen kann sie deshalb in diesem Effektgrößenbereich NICHT fungieren.

Die Auftretenswahrscheinlichkeit nur größerer ATI-Effekte ( $w^2 \geq .20$ ) kann inzwischen als gering, ihre Erwartung als eher

unrealistisch eingeschätzt werden; Über die Bedingungsstruktur kleinerer bis mittlerer ATI-Effekte ( $w^2 \approx .20$ ) hingegen sind bei der unzureichenden Präzision der meisten dazu vorgelegten Untersuchungen keine verlässlichen Aussagen zu treffen. Die Prüfbedingungen künftiger ATI-Arbeiten sind deshalb vor allem auf die Anhebung ihrer Teststärke 1- $\beta$  und die Senkung der  $\beta$ -Fehlerwahrscheinlichkeit auszurichten. Aus der Bedingungsstruktur dieser Größen (vgl. COHEN 1969, TREIBER 1976) ergeben sich verschiedene Optimierungsmöglichkeiten, die weiter unten aufgeführt werden (4.).

### 3.3 ATI-ERKLÄRUNG

Die Suche nach und Entdeckung von Merkmal-Methoden-Interaktionen umschreibt derzeit noch die wichtigste Problemstellung des ATI-Programms; die Erklärung ihres Zustandekommens zählt dazu nicht. ATI-Befunde werden denn auch bevorzugt nur zur Problematisierung übergeneralisierter Prognosen und Erklärungen des (schulischen) Lernens und deren nonadaptiver Technologieanwendung in (schulischen) Lehr-Lern-Systemen herangezogen; ihre eigene Erklärungsbedürftigkeit wird indes selten thematisiert, ihr Auftreten zumeist durch post-hoc-Überlegungen (vgl. dazu PEUCKERT 1974) plausibel zu machen versucht. Systematische Versuche zur Überbrückung des Erklärungsabstandes zwischen unabhängigen Methoden- und Lernermerkmalen und abhängigen (Lern)Variablen unter expliziter Nennung der dabei herangezogenen explanativen Konstrukte, theoretischen Verknüpfungen und der sie stützenden empirischen Evidenzen fehlen hingegen. Diese Voraussetzungen für eine erfolgreiche Erklärung von ATI-Effekten sind aber auch beim gegenwärtigen Explikations- und Überprüfungsstand der das Erreichen eines Lehrziels (Lehrzielkomplexes) bedingenden Instruktionsvarianten und der ihren Einfluß modifizierenden Lernermerkmale nicht gegeben.

Vielmehr verdeutlichen erst Bemühungen, diesen Erklärungsabstand abzubauen,

- daß die Reduktion der hypothetischen Bedingungsstruktur von

Lernprozessen und Lehreffekten auf (in der Regel) nur zwei Klassen von Einflußquellen (Lehrmethoden, Lernermerkmale) eine extreme Vereinfachung bedeutet und für die Konstruktion von geeigneten Prüfbedingungen eine so weitgehende Kontrolle konkurrierender Einflüsse impliziert, wie sie nur unter idealisierten Testbedingungen zu sichern sein dürfte; (= mangelnde Generalisierbarkeit)

- daß eine Rekonstruktion des Zustandekommens von Lehreffekten im Sinne von ATI-Effekten folglich nur einen vergleichsweise schmalen Ausschnitt ihres hypothetischen Bedingungsgefüges erfaßt, da ATI-Effekte ihrerseits als kontext-, situations-, aufgaben- sowie lerner- und lehrerspezifisch zu denken sind; (= highly niche-ATI's)
- daß folglich systematische Explikations- und Erklärungsversuche von ATI-Befunden zunächst nur zu einer Ausweitung und Vertiefung ihres Problembereichs führen und von daher eine hochgradige Bereichsspezifität und -selektivität von theoretischen Problemlösungen erwarten lassen;
- daß schließlich das ATI-Programm auch auf der Erklärungsebene vorhandene Parzellierungstendenzen in der pädagogisch-psychologischen Forschungspraxis akzentuiert wie akzentuiert.

### 3.4 ATI-ANWENDUNG

Das Auftreten eines praktisch bedeutsamen ATI-Effektes kann in verschiedenen Anwendungsbereichen ebenso erwünscht wie unerwünscht sein und somit, je nach Beurteilung, eine der folgenden beiden Nutzungsformen nahelegen:

- (1) Erwünschte Merkmals-Methoden-Kombinationen werden übernommen und bilden die Grundlage eines schüleradaptiven Lehr-Lern-Systems (durch innere oder äußere Differenzierung);
- (2) Unerwünschte Interaktionseffekte werden in ihrer Auftretenswahrscheinlichkeit in einer realen Anwendungssituation durch Reduzierung der Inter- und Intravarianz von Lernumweltsbedingungen und/oder durch Minimierung interindividueller Lernermerkmalsdifferenzen soweit zu verringern versucht, daß sie die anschließende Realisierung einer für alle einheitlichen Unterrichtsmethode ermöglichen (vgl. SIMONS 1973, HALLER 1975).

Unabhängig von der Klärung der präskriptiv-normativen Zielproblematik, welche dieser Konsequenzen in welchem Beurteilungssystem jeweils präferabel erscheint, setzt aber die Anwendbarkeit von ATI-Befunden nach dem Argument "Sollen impliziert Können" den Nachweis voraus, daß durch ihre Kenntnis die einem Instruktionssystem zur Verfügung stehenden Ressourcen zur Erreichung gesetzter Ziele besser als bisher erreicht werden können. Ansätze zu entsprechenden inter- oder intrasystemvergleichenden Analysen sind jedoch nicht vorhanden (vgl. RHETTS 1972).

Ein Effektivitätsindikator eines adaptiven Instruktionssystems ist aber zumindest in Maßen der praktischen Bedeutsamkeit des durch ATI-Effekte erklärten Anteils an der Gesamtvarianz in abhängigen Variablen zu sehen. Zieht man dabei die in der Teststärkeanalyse von TREIBER (1976) berichteten Ergebnisse als Schätzungen für die Größe der zu erwartenden ATI-Effekte heran, so zeigt sich, daß schon unter mehr oder minder schulfernen (also hinreichend kontrollierbaren und deshalb günstigen) Prüfbedingungen Interaktionen mit einer Effektgröße  $w^2 \geq .20$  nicht zu entdecken waren; es erscheint deshalb realistisch, allenfalls mit Wechselwirkungen kleinerer Effektgröße zu rechnen und den inkrementalen Nutzen von ATI-Befunden für schulische Lehr-Lern-Situationen eher bescheiden einzuschätzen.

ATI-Untersuchungen beschrieben bisher jedoch nur die Effektivität alternativer Unterrichtsmethoden bei fixiertem Kostenniveau. Bei Berücksichtigung der für verschiedene Unterrichtsmethoden aufzubringenden Gesamtkosten ließen sich möglicherweise aber bisher nur ordinale in disordinale Interaktionen transformieren, wenn als abhängige Variable nicht mehr der (je nach Unterrichtsmethode variierende) Lernerfolg erhoben werden würde, sondern der (z.B. monetär beschreibbare) Nettogewinn aus dem mit alternativen Unterrichtsmethoden zu erzielende Gesamtnutzen abzüglich der dafür aufzu-

bringenden Gesamtkosten. Dies wurde in Abb. 1 mit der Para-

Hier Abb. 1 einfügen

llelverschiebung der ursprünglichen (d.h. ohne Kostenberücksichtigung erstellten) Regressionslinie für (die zwar effektivere, aber auch kostenintensivere) Methode A zum (kostenkorrigierten) A' verdeutlicht.

Es wird in Kosten-Wirksamkeits-Analysen ferner noch zu klären sein,

- wie der durch ein adaptives Instruktionssystem zu erzielende (individuelle und institutionelle) Gewinn gegenüber dem eines fixen Systems zu bewerten ist,
- mit welchen Kosten die Einführung eines adaptiven Instruktionssystems gegenüber einem herkömmlichen fixen System verbunden ist;
- welches Wirksamkeits-Kosten-Verhältnis beide Instruktionssysteme aufweisen,

um schließlich angeben zu können, welches dieser Systeme bei festgelegtem Mindestwirksamkeits- und Maximalkostenniveau (vorläufig) vorzuziehen ist (vgl. WEISS 1976).

Ohne Klärung der ATI-spezifischen Kosten-Wirksamkeits-Relationen reduziert sich die Anwendungskonsequenz der bisher vorgelegten Befunde aber zunächst auf die Delegitimation und Destruktion der Grundlagen monomethodischer Unterrichtsmodelle und lernermerkmalszentrierter Selektions- und Klassifikationsformen. An deren Stelle tritt vielmehr zunehmend die (wenn auch abstrakte) Einsicht, deren Handhabung auf eine nach Schüler(gruppen)differierende bzw. nach Lernumweltbedingungen variierende Weise vornehmen zu müssen, ohne jedoch positiv die Realisationsmodi dafür angeben zu können. Konsequenterweise erscheint es bei dieser Sachlage denn auch, die Organisationsform heterogener Lerngruppen zugunsten einer (mehr oder minder) vollständigen Individual- oder Selbstinstruktion aufzugeben. Individualisierungs- und Differen-



zierungsbemühungen in heterogenen Lerngruppen sehen sich jedoch in dem prekären Dilemma, nach Falsifikation der üblichen Homogenitäts-, Stabilitäts- und Uniformitätsannahmen weder auf die herkömmliche starre Organisation von Lehr-Lern-Situationen zurückgreifen zu können noch die selbstverbreiteten Ansprüche oder die an sie gestellten Erwartungen auf Flexibilität und Adaptivität befriedigen oder andererseits offen frustrieren zu können. Entsprechende curriculare Empfehlungen reduzieren sich denn auch zumeist auf den formalen Hinweis (vgl. FLAMMER 1975a, RADDATZ 1975), dann eben - schon zur Minderung von Unterlassungsfehlern und Irrtumsriskiken - mit möglichst vielen Unterrichtsvarianten zu experimentieren (Mehrwegmethode) und Entscheidungen über die Zuordnung von Schülern zu alternativen Unterrichtsangeboten und Lernwegen möglichst flexibel und reversibel zu lassen. Hinweise, für welche Schüler aber welche Lehrmethoden, Lernhilfen, -materialien, -zeiten und Sozialformen wie angemessen sind, lassen sich ihnen nicht entnehmen.

Bei der Vielzahl berücksichtigenswerter Klassifikations- und Kombinationsmöglichkeiten von Schülergruppen und Instruktionsvarianten bezweifeln mehrere Autoren ohnehin, ob ATI-Befunde überhaupt für die Einrichtung adaptiver und flexibler Curricula geeignet sind, deren Innovationsaufwand die Belastungs- und Zumutbarkeitsgrenzen institutionalisierter Lehr-Lern-Systeme und die kognitive Kapazität der in ihnen Handelnden nicht überschreitet (vgl. HILLS 1971). So lassen sich z.B. gegenwärtig zumindest die folgenden Merkmals-Methoden-Kombinationen als Beispiele für ATI-Befunde mit möglichen Anwendungschancen anführen (vgl. FLAMMER 1973, 1975a; BERLINER & CAHEN 1973; LESSER 1972):

---

Hier Tab. 3 einfügen

---

Eines der zentralen Probleme bei der Nutzung positiver ATI-Evidenzen (z.B. durch Unterstützung binnendifferenzierender Instruktionsformen) dürfte deshalb die Explikation eines unterrichtspraktischen Optimalitätsbereiches sein, in dem sich unterhalb der Maximalforderung eines Individual- oder Tutorialunterrichts hinreichend flexible Handlungsmodelle trotz unvermeidbarer Komplexitätsgrenzen effektiv und effizient einsetzen lassen.

Die Weitmaschigkeit der durch ATI-Befunde bislang herstellbaren "Strukturgitters" für die inhaltliche Ausfüllung dieser Handlungsmodelle und die Erprobung ihrer Anwendungsgrenzen und -möglichkeiten erschwert oder verhindert auch die Beibehaltung der üblichen Interaktionsformen zwischen Forschungs- und Unterrichtspraxis (erstere als Innovationsproduzent, letztere als Innovationsabnehmer). Es liegen dafür nämlich keine fertigen Differenzierungs-/Individualisierungsprodukte vor, deren standardisierte Anwendung an detaillierte Anweisungen gebunden und damit begründet werden könnte, daß ihr Optimierungsnutzen bei Modifikation ihrer Organisations-, Instruktions- und Evaluationsschemata durch Lehrer, Schüler, Kontextbedingungen und Verlaufsmerkmale eines Lehr-Lern-Prozesses gefährdet wäre. Diese Beratungs-, Orientierungs- und Strukturierungsschwäche des ATI-Programms bedeutet aber zugleich eine Offenheit des Problemlösungsraums hinsichtlich der Erhöhung der Flexibilität und Adaptivität in Lehr-Lern-Situationen, für die konstruktiv die Entwicklung und Erprobung "offener" Curricula in Formen der Handlungsforschung empfohlen werden kann (vgl. HEIPCKE & MESSNER 1973, SACHS & SCHEILKE 1973).

## 3.5 ZUSAMMENFASSUNG

Das ATI-Programm gilt der Suche nach sowie der Entdeckung und Sicherung möglichst bedeutsamer Wechselwirkungen von Schülermerkmalen und Unterrichtsmethoden auf abhängige (Lern)Variablen.

- (1) Begründung und Formulierung der ATI-Problemstellung begünstigen und beschleunigen zunächst die Problematisierung herkömmlicher Versuche, Schülerlernleistungen entweder nur anhand einzelner Schülermerkmale vorherzusagen oder durch Variation von Instruktionsbedingungen zu erklären. Die dabei zugrunde gelegten Stabilitäts-, Konstanz- und Homogenitätsannahmen sind inzwischen als durch hinreichend viele ATI-Evidenzen falsifiziert zu betrachten.
- (2) Die weitere Ausarbeitung der ATI-Problemperspektive hingegen weist trotz mehrere Befunde, die als prototypische Beispiele von Anwendungsmöglichkeiten ihrer Kernannahmen gelten können (vgl. TALLMADGE & SHEARER 1969), ein derzeit stagnierendes Entwicklungsmuster auf; hinreichend viele gut bewährte ATI-Beispiele mit Generalisierbarkeitschancen auch auf außereperimentelle Lehr-Lern-Situationen liegen jedenfalls nicht vor. (vgl. BIGGS 1976, FLAMMER 1975ab).
- (3) Zur Erklärung dieses Sachverhaltes wurde eine Reanalyse des bisherigen Verlaufs des ATI-Programms versucht. Diese führte zu folgenden Ergebnissen:
  - (a) Die **G e n e r i e r u n g** gehaltvoller ATI-Hypothesen wurde bislang vor allem durch empiristische Suchstrategien auf der Basis unsystematischer Plausibilitätsannahmen zu organisieren versucht. Bei der Vielzahl denkbarer Merkmal-Methoden-Kombinationen und der Selektions- und Strukturierungsschwäche theoriearmer bis -loser Evidenzüberlegungen ließ sich die Entdeckung von Interaktionseffekten nur unerheblich verbessern.
  - (b) Die **V e r k n ü p f u n g** selemente von ATI-Annahmen

(Schülermerkmale, Unterrichtsmethoden, Kriteriumsvariablen) weisen Mängel in ihrer **Explikation** und **Operationalisierung** auf, welche die **Entdeckungschance** replizierbarer ATI-Befunde **zusätzlich einschränken**.

- (c) Die **Ü b e r p r ü f u n g** von ATI-Annahmen bedient sich noch eindeutig suboptimaler **methodisch-statistischer Analysepläne** und **Testprozeduren** und erfolgt unter konservativ-restriktiven Prüfbedingungen, welche die Entdeckung von **Interaktionseffekten kleinerer bis mittlerer Größe**, wie sie in der **Regel** zu unterstellen sind, eher unwahrscheinlich **machen** und die Beibehaltung der Nullhypothese **erheblich erleichtern**.
  - (d) Die systematische **E r k l ä r u n g** des Auftretens von ATI-Effekten, die nicht erst **post factum** erfolgt, scheitert - wie in (a) - an der **Absenz geeigneter Konstrukt- und Theorie-segmente**, die **einen** differentiellen Verlauf und Effekt eines **Lernprozesses** in Abhängigkeit von **internen Merkmalen des Lernenden** und **externen Merkmalen seiner Lernumwelt** erwarten lassen.
  - (e) Die **N u t z u n g** von ATI-Befunden **in natürlichen** (z.B. schulischen) **Lehr-Lern-Situationen** ist an die **Bearbeitung, Klärung und Lösung mehrerer Kosten-Wirkungs-Probleme** gebunden, die **derzeit noch ausstehen**. Der bescheidene **Optimierungszuwachs**, wie er an der **Effektgröße** von bislang **berichteten Merkmals-Methoden-Interaktionen abzulesen ist**, sowie die Existenz verschiedener (**individueller und institutioneller**) **Rezeptionsbarrieren räumen ihnen ohnehin eine nur geringe Anwendungschance ein**.
- (4) Die Rekonstruktion dieser **Mängeldimensionen** in vorliegenden ATI-Arbeiten legt zunächst den **Schluß** nahe, die **überzufällig häufige Ablehnung von ATI-Annahmen** nicht auch schon als **validen Indikator für die Nicht-Existenz**

entdeckenswerter Interaktionseffekte und die daraus folgende zwangsläufige Stagnation in der künftigen Entwicklung dieses Forschungsprogramms zu werten. Die nachweisbare Abhängigkeit eines Interaktionseffektes von situations-, kontext-, populations- und aufgabenspezifischen Bedingungen, von der Güte der verwendeten Erhebungsinstrumente, der Konsistenz der realisierten Instruktionsvarianten, der Kontrolle störender Einflüsse sowie der Art der herangezogenen Prüfprozeduren und -kriterien bilden den Plausibilitätshintergrund entsprechender exhaustiver Argumente. Da bislang nur von einer suboptimalen Realisierung des ATI-Programmauftrags die Rede sein kann und Art und Ausmaß der Auswirkungen der zuvor genannten Realisationsdefizite kaum abzuschätzen sind, weisen sämtliche Versuche, den inkrementalen Erklärungs- und Vorhersagewert der ATI-Perspektive für individuelle Lernleistungen hinreichend eindeutig einzuschätzen, ein vergleichsweise hohes Irrtumsrisiko auf. Es erscheint deshalb sinnvoll, sich der Beantwortung dieser Frage erst nach Behebung der genannten Mängel wieder zuzuwenden. Insbesondere erscheinen denn auch jene Argumente verfrüht, Ausarbeitung und Erweiterung des ATI-Programms seien als offizielles Eingeständnis des Scheiterns einer konsequenten Nomologisierungstrategie in der Psychologie zu werten. Die damit anerkannte Parzellierung des jeweiligen Objektbereichs und die ihr entsprechende Abnahme des Informationsgehalts nomologischer Hypothesen beschleunige im Zuge der Programmanwendung nur noch eine degenerative Entwicklung, an deren Ende schließlich "ein chaotisches Durcheinander relativ simpler, an-ändernder, trivialer und langweiliger Minimalhypothesen" stehe (MÜNCH 1973: 172) (vgl. auch DICK 1972, HOLZKAMP 1972, LAKATOS 1970:176, MERTENS 1975). Die Gültigkeit

dieser lokalen Stagnations- und Degenerationserwartungen für den künftigen ATI-Erkenntnisfortschritt kann sich zwar auf den bisherigen Verlauf dieses **Forschungsprogramms** stützen (und erscheint insofern auch **nicht** kontrafaktisch). Doch sind dessen Leistungsdefizite auch durch die z.T. erheblichen Realisationsschwächen bei der Formulierung und Präzisierung von ATI-Annahmen und **der** Planung, Durchführung und Auswertung entsprechender **Prüfversuche** erklärbar, die einer Verbesserung durchaus **zugänglich** sind und damit zu einer Korrektur der bisherigen Entwicklungslinie des ATI-Forschungsprogramms beitragen können. Solange Art und Ausmaß des Optimierungseffektes dieser **forschungspraktischen** (vor allem **methodisch-explikativen**) **Steuerungsgrößen** unbekannt sind, erscheinen weiterreichende Entscheidungen z.B. über die prospektive Adäquanz einer **Nomologisierungstrategie** in diesem Gegenstandsbereich **verfrüht**. Als Argumentationshilfe sind die Rekonstruktionsergebnisse des ATI-Programms jedenfalls (derzeit noch), da sie **auch** mit anderen z.B. **forschungspraktische Erklärungen** **verfeinbar** und demnach **unspezifisch** sind, unzureichend.

#### OPTIMIERUNGSANSÄTZE

Hinweise auf und Vorschläge für die Nutzung dieser **forschungspraktischen Optimierungsmöglichkeiten** beziehen sich dabei auf die folgenden **Mängeldimensionen**:

- (1) Die **Generierung** von **ATI-Annahmen** erweist sich, nach dem relativen Mißerfolg **empiristischer Heuristikstrategien**, als zunehmend **abhängig** von einem Explikations- und Bewährungsgrad von **Konstrukten** und Theorien unterschiedlicher Lehr-Lern-Prozesse, der erst noch erreicht werden muß. Dazu bieten sich verschiedene **Heuristikmodelle** und **Konzeptualisierungsschemata** an, die von RHETTS (1972, 1974), SALOMON (1972) und WEBSTER & MENDRO (1974) beschrieben wurden. Sie empfehlen, die Formulierung von ATI-Aussagen **abhängig zu machen**

- (a) von einer sachstrukturellen Analyse der das Lernkriterium umschreibenden Aufgabenmenge und ihrer Teilkomponenten, vor allem sofern diese eine sequentielle oder hierarchische Beziehung aufweisen;
- (b) von einer adäquaten Beschreibung der für die Lösung dieser Aufgabenmenge notwendigen (Vor)Kenntnisse und Fähigkeiten, der sie gleichfalls beeinflussenden Schülerinteressen und -präferenzen sowie der zum Aufbau, Erwerb und Wandel dieser Merkmale führenden Lernprozesse;
- (c) von einer begründeten Festlegung der (z.B. kompensatorischen vs. medialen) Funktion einer schülerbezogenen Adaption des Instruktionsangebots;
- (d) von einer darauf bezogenen Konstruktion alternativer Unterrichtsmethoden unter Berücksichtigung von Art, Ausmaß und Richtung der Steuerungsfunktion und -kompetenz der unterrichtenden Lehrperson.

Der Erfolg dieses Vorgehens kann an folgendem Beispiel verdeutlicht werden:

- (a) ROHWER und seine Mitarbeiter verwendeten als bevorzugte abhängige Variable ihres Arbeitsprogramms die (reproduktive oder rekognitive) Behaltensleistung in experimentellen Paarassoziationsaufgaben, deren Teilkomponenten und Binnenstruktur vergleichsweise gut bekannt waren (vgl. MURDOCK 1974) und in weiterem Verlauf ihrer Untersuchungen noch präzisiert werden konnten (vgl. ROHWER 1974).
- (b) Davon ausgehend wurden Extremgruppen mit besonders guter vs. schlechter Behaltensleistung kontrastiert und im Anschluß an diese Lernaufgabe nach der Art ihres Vorgehens bei der Bearbeitung dieser Aufgabe befragt (vgl. BUGELSKI 1962, MARTIN, BOERSMA & COX 1965, RUNQUIST & FARLEY 1964). Aus den Angaben der erfolgreich Lernenden ließ sich dann als bevorzugte elaborative Lösungstaktik rekonstruieren, die nahezu sinnlosen Elemente des paarweise dargebotenen Lernmaterials so zu transformieren, daß diese sich in einen (in der Regel einfachen S-V-O-Satz) oder in eine entsprechende mental vorstellbare Bildstruktur integrieren lassen (vgl. MONTAGUE 1972).
- (c) Diese Angaben wurden dann als hypothetische Bedingungen des paarassoziativen Lernerfolgs betrachtet und für die Konstruktion und Realisierung von solchen Lernmaterial-

bzw. Instruktionsvarianten verwendet, in denen erfolglos lernenden Vpn die fehlenden Transformations-, Elaborations- und Integrationsformen vom V1 bereitgestellt werden oder in denen deren Verwendung fremdinduziert wird, und zwar jeweils in Übereinstimmung mit oder im Kontrast zu der präexperimentell ermittelten Ausprägung ihrer visuell-imaginativen vs. verbal-auditiven Informationsverarbeitungspräferenz/kapazität (vgl. zusammenfassend TREIBER & GROEBEN 1976).

- (d) Als Ergebnis zeigte sich, daß die verbale oder visuelle Fremd- oder Eigenelaboration der zu lernenden Paarassoziationen den Lernerfolg in Abhängigkeit von der interindividuell variierenden Disposition bestimmt, Lernmaterial verbal oder visuell zu kodieren, zu verarbeiten, zu speichern oder abzurufen (LEVIN et al. 1971, MALLORY 1972, ROHWER & LEVIN 1971). Ordinale bis disordinale Interaktionseffekte zwischen dem Lernermerkmal "Elaborationspräferenz" und dem Instruktionsfaktor "Elaborationsmodalität" berichten MALLORY (1972) und TREIBER (1974).

Die Generierung von ATI-Aussagen ist darüberhinaus aber vor allem auf die systematisierte Stimulierung, Förderung und Regelung der Theorienbildung in diesem Bereich angewiesen und von einer hinreichenden Kenntnis und erfolgreichen Beeinflussung ihrer Steuerungsgrößen abhängig. Metatheoretische Erklärungsversuche der Entstehung und weiteren Entwicklung von Theorien unter Berücksichtigung vor allem wissenschaftsinterner (z.B. forschungspraktischer) Eigenregulative weisen aber bislang noch erhebliche Konzeptualisierungsschwächen (z.B. bei der Bereitstellung einer einheitlichen semantischen Informationstheorie für die Beschreibung des theoretischen Entwicklungsverlaufs) und elementare Problemlösungslücken auf, die erst noch über die Erprobung konkurrierender Theorien des Erkenntnisfortschritts zu lösen wären (vgl. SPINNER 1974) und die deshalb die Erstellung eines Programms positiver Heuristiken (sensu LAKATOS 1970) mit eindeutig innovationsfördernden Konsequenzen derzeit unrealistisch machen.

Dennoch lassen sich im bisherigen Programmverlauf zumindest fortschrittshemmende Komponenten identifizieren und, soweit hinreichend bekannt, auch eliminieren.

Dies gilt vor allem für die Restriktivität der ATI-Prüfbedingungen: durch überzogene Signifikanz- und Bewährungsauflagen wurden nämlich Artikulation und Präzisierung von ATI-Annahmen, die beim defizitären Stand des theoretischen Hintergrundwissens über Lehr-Lern-Prozesse zwangsläufig unterentwickelt, nur bedingt überprüfbar und deshalb auch erklärungs-schwach sind, nicht etwa gefördert, sondern erheblich behindert. Insofern sind diese Prüfkriterien auch zu korrigieren, da sie, offenbar auch in Unkenntnis der Bedingungsstruktur der Teststärke empirischer Untersuchungen, eine nur virtuelle Falsifikationsinstanz für ATI-Annahmen über Merkmal-Methoden-Wechselwirkungen realistischer Effektgrößen darstellen, diese Funktion aber (bei der in aller Regel viel zu hohen  $\beta$ -Fehlerwahrscheinlichkeit) gar nicht übernehmen können. In der aktuellen Explorationshypothese dieses Programms bedarf es demnach solcher Analysen, die auch für die Entdeckung von Interaktionen von mittlerer oder kleiner Effektgröße hinreichend präzise sind bzw. die Annahme der sie falsifizierenden Nullhypothese bei bekanntem und akzeptablem kleinem Fehlerisiko ermöglichen. Dieses Überprüfungs-niveau ist bislang aber in der Regel gar nicht erreicht worden.

Darüberhinaus unterliegen ATI-Annahmen aber auch bei für sie negativen Testausgängen keinem automatischen Falsifikations- und Eliminationszwang, sondern bedürfen dazu pragmatischer Zusatzkriterien, die solche Entscheidungen hinsichtlich ihres Beitrags für den Erkenntniszuwachs zu bewerten haben. Die Beurteilung einer Prüfungshypothese nur anhand ihres Bestehens oder Nicht-Bestehens in empirischen Prüfversuchen würde die tatsächliche normative-regulative Leistungsfähigkeit solcher Methodologiekriterien, -regeln und -standards erheblich überschätzen, den Überprüfungs-vorgang vor allem bei explorativen ATI-Annahmen unzulässig vereinfachen und beim nachweislich

ungenügenden Stand des testrelevanten Beobachtungs- oder Hintergrundwissens auch zum allzu voreiligen Ausschluß entwicklungsbedürftiger und -fähiger ATI-Theorieansätze führen.

Die verbesserte Generierung von ATI-Hypothesen ist folglich zumindest an die Liberalisierung/Beseitigung von die Theorieentwicklung und -entstehung hemmenden Einflüssen gebunden, zu denen konservative Prüfstrukturen bisheriger Untersuchungen ebenso zählen wie die naiv-falsifikationistische Überschätzung der Entscheidungsrelevanz empirischer Testausgänge beim durchweg problematisierbaren Stand des relevanten Hintergrundwissens.

Dazu kann aber auch eine veränderte Problemperspektive und Suchrichtung wie auch ein multistrategisches Vorgehen bei der empirischen Analyse von Lehr-Lern-Situationen beitragen. Als Datenquellen bieten sich dabei z.B. die Ergebnisse von Multivariaten Interdependenzstudien an, in deren Mittelpunkt z.B. die reflexiv-subjektive Erklärung von Lehrern ihres eigenen differentiellen Verhaltens gegenüber unterschiedlichen Schülergruppen (vgl. dazu BROPHY & GOOD 1974) steht und welche die extensive Nutzung der verfügbaren Such-, Auswertungs- und Variablenreduktionsverfahren beinhalten; von laborexperimentell streng kontrollierbaren Prüfversuchen, die eine möglichst eindeutige Überprüfung der Parameterstruktur von voraussetzungsarmen Lernprozessen in formalisierten Modellen ermöglichen; aber auch von naturalistischen bis manipulativen Einzelfallstudien von Lernenden unter variierenden Instruktionsbedingungen. Da bislang in der Regel nur Lernermerkmale mit vergleichsweise hoher zeitlicher Konstanz untersucht wurden, können diese Analysetechniken auch die bislang kaum thematisierte Möglichkeit einer bei i n d i v i d u e l l e n Lernermerkmals-Veränderungen eintretenden Interaktion mit (konstanten oder variierenden) Lern-

oder Unterrichtsbedingungen überprüfen helfen. Vor allem bei zeitlicher Instabilität der jeweiligen Merkmalsausprägung empfehlen sich deshalb Längsschnitt- und Einzelfallanalysen.

Einzelfallanalysen wurden bisher vor allem für die Erfassung von (Psycho)Therapieeffekten genutzt (vgl. DAHME 1975, HUBER 1973) und lassen sich methodisch wie folgt unterscheiden (vgl. ausführlich PETERMANN & KNOPF):

Einzelfallstudien (Single-subject-designs) (vgl. SHINE & BOWER 1971, SHINE 1973ab, 1974, 1975); Anwendungen der P-Technik (vgl. CATTELL 1946, 1951 ab) und ihrer (z.B. zeitkorrigierten) Erweiterungen (zur "time-corrected lead and lag-technique" vgl. CATTELL 1966); Zeitreihenanalysen (HOLTZMAN 1963, HUBER 1967, 1973, BALDWIN 1950).

Das Ziel dieser Vorgehensweisen ist dabei generell die Ausarbeitung einer Klasse von Theorien des (schulischen) Lernens (vgl. UNDERWOOD 1975),

- deren allgemeinspsychologische Aussagen über die Bedingungs-, Verlaufs- und Effektstruktur eines Lernprozesses einen von inzidentellen Individual/Personparametern separierbaren und möglichst wenig begrenzten Geltungsbereich aufweisen;
- deren differentialpsychologische Komponenten gleichwohl aber Art und Bedeutung von lernerfolgsrelevanten Personparametern begründen und erklären.

Probabilistische Testmodelle in ihren RASCH-Varianten scheinen für die Präzisierung, Überprüfung und Bestätigung solcher Annahmen zumindest bei einfachen und voraussetzungsarmen Lernprozessen (vgl. FISCHER 1971:381) besonders geeignet. Können zentrale Modellannahmen (wie Homogenität der Testaufgaben und vor allem spezifische Objektivität) nämlich beibehalten werden, liesse sich dadurch die durch eine Vielzahl immer neu eingeführter Personparametern zunehmende beschränkte Gültigkeit allgemeiner Strukturaussagen schrittweise wiederherstellen und ausweiten. Auf diese Möglichkeit, probabilistische Ansätze in der Testtheorie für die Konstruktion solcher Beobachtungsannahmen und Meßinstrumente zu nutzen, die eine von inzidentellen

Individualparametern zumindest separierbare und verallgemeinerungsfähige Überprüfung allgemeiner Strukturaussagen (z.B. über den Instruktionseffekt unterschiedlicher Unterrichtsmaßnahmen als Ursache für die Veränderung von Itemschwierigkeiten) zulassen, hat vor allem FISCHER (1974, Kap. 16) hingewiesen. MÖBUS & SIMONS (1975) wandten diesen Ansatz für eine probabilistische Explikation des Konzeptes der Testfairness an, das eine strukturelle Analogie zur ATI-Perspektive aufweist und das deshalb im Zuge seiner Ausarbeitung (vor allem hinsichtlich der Konstruierbarkeit eines gegenüber ethnischen und sozialen Gruppen fairen und psychologisch relevanten Tests) auch die Adäquanz eines RASCH-orientierten Ansatzes im Rahmen des ATI-Programms zu beurteilen erlauben sollte.

Andererseits kann sich dabei auch die auf Merkmals-Methoden-Wechselwirkungen eingeeengte Analyseperspektive als zu eng erweisen für eine hinreichend genaue Beschreibung, Erklärung, Vorhersage und Beeinflussung von Lehr-Lern-Prozessen. Das ATI-Programm ist in dieser Hinsicht möglicherweise nur als intermediärer Versuch aufzufassen, Anomalien seiner explanativen wie prognostischen Vorläufer aufzufangen, Schülerleistungen entweder nur durch Schülermerkmale vorherzusagen oder nur durch Veränderung seiner instruktionsspezifischen Lernumwelt zu erklären (vgl. HUNT 1975). Die Ausarbeitung dieser interaktiven Perspektive trägt zwar einerseits zu einer Problematisierung, Ablösung und Verdrängung von deren Konstanz- und Homogenitätsannahmen bei, die sich als unrealistisch erwiesen. Andererseits ist sie selbst transformierbar in eine Problemstellung, in der sich die gesamte Parameterstruktur von Lehr-Lern-Prozessen als aufklärungsbedürftig und für die sich eine selektive Thematisierung des Einflusses nur von zwei Variablenklassen (Lernermerkmale, Lehrmethoden) als unergiebig, da zu voraussetzungsvoll erweist. Insofern akzentuieren Leistungsdefizite und Ent-

wicklungsstagnation des ATI-Programms auch zentrale Problemlücken der aktuellen Lehr-Lern-Forschung, von deren Schließung erst eine längerfristig erfolgreiche Verbesserung seiner eigenen Problemlösungsfähigkeit zu erwarten ist.

(2) Zur Überprüfung von ATI-Annahmen

Die Generierung von ATI-Annahmen erwies sich bereits als abhängig von der verbesserten Präzision empirischer Prüfversuche und einer Liberalisierung der sich an ihren zumeist negativen Ausgang knüpfenden Falsifikations- und Eliminationszwänge. Beide Voraussetzungen berühren damit die Konstruktion von bislang inadäquat ausgestalteten und künftig auch verbesserungsfähigen Prüfbedingungen, -standards und -kriterien, die stärker als bisher auf die Ermöglichung, Förderung und Regulierung der Theoriegewinnung und -entwicklung auszurichten sind.

Dazu zählt zum einen die bessere Nutzung der verfügbaren Möglichkeiten zur Verbesserung der Präzision künftiger ATI-Studien und zur Senkung der  $\beta$ -Fehlerwahrscheinlichkeit bei der Überprüfung von Interaktionsannahmen auch in realistischen Effektgrößenbereichen. Aus der Bedingungsstruktur von  $\beta$  und  $1-\beta$  (vgl. BREDEKAMP 1972, COHEN 1969) ergeben sich diese Möglichkeiten wie folgt:

(1) Erhöhung des Stichprobenumfangs

Diese Variable ist allerdings wegen der erheblichen Stichprobengröße, wie sie für die Entdeckung kleinerer Interaktionseffekte in varianzanalytischen Designs nötig ist, nur bedingt zu beeinflussen (vgl. COHEN 1969: 368).

BORICH & GODBOUT (1974) zeigen aber,

- daß die unabhängige(n) Lernervariable(n) in den meisten bisher vorgelegten ATI-Arbeiten als zweistufige diskrete Variable(n) behandelt wurde(n), wobei der Median üblicherweise als Klassifikationswert gilt;
- daß darüberhinaus aber die Bildung von Extremgruppen aus der (den) unabhängigen Lernervariable(n) sowohl eine Präzisionserhöhung ermöglicht als auch die Kosten eines erweiterten Gesamtstichprobenumfangs  $N$  relativ zu denen einer

einfachen Dichotomisierung zu limitieren erlaubt. Sie erstellten dafür eine Übersichtstabelle, der die für ein gewünschtes Präzisionsniveau ( $1-\beta$ ) optimale Anzahl von Elementen zweier Extremgruppen mit unterschiedlichen Perzentil-Werten zu entnehmen ist. Die für die Bildung von Extremgruppen erforderliche Stichprobenerweiterung erweist sich bei konstantem Präzisionsniveau  $1-\beta$  als weitaus weniger aufwendig als die Erhöhung von  $N$  im Rahmen der üblichen Gruppierung von Versuchspersonen (z.B. über den Median der unabhängigen Lernervariable).

(2) Verstärkte Kontrolle der Fehlervarianz (= within-SS)

Die Wahrscheinlichkeit  $1-\beta$  entspricht der Präzision (Power) einer Untersuchung, mit der diese einen Effekt zu entdecken vermag. Sie hängt u.a. von der Größe des Versuchsfehlers (Testvarianz) ab, der durch experimentelle (z.B. Randomisierung, Parallelisierung, Einführung zusätzlicher Faktoren etc.) und/oder statistische Kontrolltechniken (z.B. Kovarianzanalysen) reduziert werden kann. Andererseits ergibt sich aus der strukturellen Gegenläufigkeit von interner und externer Validität von Prüfversuchen, daß bei erhöhter Experimentalkontrolle auch die Generalisierungsmöglichkeit auf andere (vor allem außerexperimentelle) Anwendungssituationen schwindet, so daß sich die Nutzungschance für Ergebnisse aus besser kontrollierten ATI-Untersuchungen verringert und sich die Kosten-Nutzen-Differenz eher konstant bilanziert.

(3) Anhebung des Signifikanzkriteriums  $\alpha$

Die Festlegung eines zu niedrigen Signifikanzkriteriums (z.B.  $\alpha = .01$  oder  $.001$ ) begünstigt die konservative Beibehaltung der Nullhypothese und, da sich  $\beta$  umgekehrt proportional zu  $\alpha$  verhält, die fälschliche Ablehnung ihrer Alternativhypothese. Dies wirkt sich aber gerade in heuristisch-explorativen Phasen eines Forschungsprogramms nachträglich für dessen Entwicklung aus. Falsifikations- und Eliminationsentscheidungen sind nämlich zu diesem Zeitpunkt, da eine hinreichend lange Erfolgsbilanz noch nicht vorliegt, in jedem Falle verfrüht, da über die fälschliche Beibehaltung oder Ablehnung einer ATI-Annahme zu diesem Zeitpunkt noch zu wenig Indika-

toren vorliegen. Konventionalistische Exhaustionsversuche sind deshalb auch bei der Entwicklung des ATI-Programms eher angebracht als eine konsequente Eliminationsstrategie, die Folgen einer fälschlichen Beibehaltung einer ATI-Annahme (bei relativ großem  $\alpha$ ) als weniger nachteilig einzuschätzen als die ihrer verfrühten Ablehnung (bei hohem  $\beta$ ). Deshalb kann auch die Erhöhung des bisherigen Signifikanzkriteriums von  $\alpha = .05$  oder  $.01$  auf  $\alpha = .10$  empfohlen werden.

(4) Erhöhung der Anzahl der Freiheitsgrade  $df$  für den Interaktionsausdruck.

Die Anzahl der Freiheitsgrade  $df$  für den kritischen Interaktionsausdruck hängt von der Anzahl der Bedingungsstufen von Methoden- und Schülermerkmalen ab und kann durch deren Erweiterung erhöht werden. In der Regel erfolgt aber eine Variation des Faktors "Unterrichtsmethode" schon aus pragmatischen (Kosten)Gründen über den Vergleich einer Experimental- und Kontrollgruppe, so daß sich eine Erhöhung der Anzahl der Freiheitsgrade am ehesten noch über eine Ausdifferenzierung, der gleichfalls zumeist nur dichotomisierten Persönlichkeitsvariable(n) erreichen läßt, an deren Ende eine kontinuierliche Variablenexplikation und regressionsanalytische Datenauswertung steht. Diese Vorgehensweise wird denn auch von mehreren Autoren empfohlen (vgl. BRACHT 1969, 1970; BREDEKAMP 1975; FLAMMER 1973, PLOMP 1974, SCHWARZER & STEINHAGEN 1975) und in neueren ATI-Arbeiten häufiger angewandt. Ausführlich beschreiben KERLINGER & PEDHAZUR (1973) die Anwendung und illustrieren sie an einem Rechenbeispiel.

Generell scheint es aber eines der zentralen Probleme bei der Überprüfung von ATI-Annahmen in einem theoretisch noch wenig strukturierten explorativen Stadium zu sein, eine möglichst hohe Sensibilität (Präzision) für die Entdeckung von Merkmal-Methoden-Interaktionen zu erreichen. Es liegt deshalb nahe, die bisher verwendeten Varianz- und regressionsanalytischen Verfahren für die Entdeckung von Interaktionseffekten

nicht nur durch Erhöhung der Anzahl ihrer unabhängigen Lernvariablen zu erweitern; vielmehr kann deren Interaktion im Rahmen einer mehrfaktoriellen multivariaten Varianzanalyse (MANOVA) simultan auch für mehrere abhängige Variablen überprüft werden. Dieses Vorgehen ist vor allem gegenüber dem separaten Abtesten von ATI-Hypothesen für jede einzelne abhängige Variable vorzuziehen, da sich dadurch das  $\alpha$ -Fehlerrisiko erhöht, und zwar in Abhängigkeit von der Anzahl der Testungen  $n$  wie folgt:  $1 - (1 - \alpha)^n$  (vgl. KIRK 1968: 83; sowie vor allem BREDEKAMP 1972: 114-116, der auch auf das Problem der -zumeist unbekannt- Teststärke multivariater Tests eingeht).

In der Untersuchung von ANDERSON (1949) wurde beispielsweise für jede der insgesamt 12 abhängigen Variablen ein Einzeltest durchgeführt. Somit war die Wahrscheinlichkeit, daß wenigstens ein F-Test eine signifikante Interaktion per Zufall anzeigte, nicht mehr  $\alpha = .05$ , sondern  $1 - (.9)^{12} = .49$ . Bei jedem zweiten Test war folglich die Entdeckung einer signifikanten Interaktion per Zufall schon zu erwarten.

Analog zur univariaten Varianzanalyse (ANOVA) schließen sich der Entdeckung eines signifikanten Interaktionseffektes in einer MANOVA Analysen der einfachen Haupteffekte an, wobei lineare Diskriminanzfunktionen herangezogen werden können (vgl. WOODWARD & OVERALL 1975). Interpretationen von MANOVA-Ergebnissen sind allerdings an die Existenz gut ausformulierter Hypothesen gebunden.

In Abb. 2 (I-IV) wurden die wichtigsten Prüfverfahren für ATI-Annahmen in einem integrierten Auswertungsschema zusammengefaßt (vgl. BRACHT 1970, BREDEKAMP 1975, PLOMP 1974), dem sich das für unterschiedliche Versuchspläne und Analyseverfahren geeignetste Vorgehen entnehmen läßt.

Hier Abb. 2 einfügen

Die Lern- und Entwicklungsfähigkeit eines Forschungsprogramms hängt schließlich auch von der Verknüpfung von Ergebnissen einzelner Untersuchungen untereinander ab. Diese Möglichkeit, die Aussagen früherer Forschungsergebnisse in Rekonstruktions- oder Replikationsstudien oder über iterative Analysemethoden erneut zu überprüfen, zu kritisieren und zu verbessern, wird



aber nur selten (vgl. z.B. CRONBACH & WEBB 1975) genutzt.

- BORICH (1975) zeigt beispielsweise,
- daß in Untersuchungen oft schon das Auftreten von Interaktionseffekten als unerwünscht gilt und der konfundierende (aber zugleich auch zu ATIs führende) Einfluß von Lernermerkmalen auf davon abhängige Variablen z.B. durch Kovarianzanalysen zu kontrollieren versucht wird;
  - daß post hoc aber in diesen Fällen (zumindest bei Verwendung von Kovarianzanalysen) noch nach signifikanten Interaktionen gesucht werden kann.

Ein systematischer Optimierungsweg eröffnen aber empirische Suchstrategien für den Fall mehrerer bis zahlreicher Lernervariablen, nachdem deren Interaktion mit dem Faktor "Unterrichtsmethode" in unikriterialen (über den F-Test auf Homogenität der Regressionen) oder multikriterialen (über den  $\lambda$ -Wert einer mehrfaktoriellen MANOVA) Signifikanztests zufallskritisch gesichert wurde. Ziel dieser Optimierungsmethoden ist es, aus der Gesamtzahl der Prädiktoren (Lernermerkmale) einen kleineren Satz möglichst vorhersagekräftiger Prädiktoren auszuwählen und jene Prädiktoren zu eliminieren, die in der Folge der Prädiktoren einen nur un-erheblichen Beitrag leisten für die Vorhersage der abhängigen Variable(n). Die Anzahl der zu erhebenden Lernermerkmale kann durch dieses Vorgehen ohne (wesentliche) Minderung der Vorhersagbarkeit des (der) Kriteriums (Kriterien) mehr oder minder deutlich gesenkt werden. Dabei können die folgenden beiden Vorgehensweisen angewandt werden:

**Unikriteriale Variablenreduktion**  
Die kritische Prüfinstanz für unikriteriale ATI-Hypothesen ist, sofern regressions- und nicht varianzanalytisch getestet wird, ein Vergleich des durch die vollständige Regressionsgleichung (für Haupt- und Interaktionseffekte) vs. des durch eine um das Interaktionsglied verkürzte Gleichung erklärten Varianzanteils in der abhängigen Variablen durch einen F-Test, der wie folgt definiert ist (vgl. BREDEKAMP 1975:808; KERLINGER & PEDHAZUR 1973:251ff; MCNEMAR 1969:321):

$$F = \frac{(R_1^2 - R_2^2) (N - m_1 - 1)}{(1 - R_1^2) (m_1 - m_2)}$$

$R_1$  = Anteil an der Kriteriumsvarianz, der durch Haupt- und Interaktionseffekte maximal erklärbar ist

$R_2$  = Anteil an der Kriteriumsvarianz, der nur durch die Haupteffekte erklärt wird

$m_1$  = Anzahl der Vektoren von  $R_1$

$m_2$  = Anzahl der Vektoren von  $R_2$

$N$  = Stichprobengröße

Für  $R_1$ ,  $R_2$  und  $R_1 - R_2$  können nun schrittweise Regressionsanalysen gerechnet werden (vgl. EFROMSON 1960), mit denen sich die Anzahl der für die Vorhersage der vorgegebenen R-Werte erforderlichen Prädiktoren versuchsweise reduzieren läßt. GAENSSLEN & SCHUBÖ (1973) weisen jedoch auf verschiedene kritische Punkte (z.B. Stichprobenabhängigkeit) der Verfahren hin, die nur bei größeren Stichproben vermeidbar erscheint.

#### Multikriteriale Variablenreduktion

HAEL (1971) stellt eine Verallgemeinerung der WHERRY-DOOLITTLE-Technik für die Reduktion des Lernervariablen-Satzes einer MANOVA vor, vergleicht sie mit der von EFROMSON (1960) vorgeschlagenen schrittweisen regressionsanalytischen Variablen-selektion und geht auch auf verfahrenstechnische Probleme ein.

Einige dieser Vorschläge eröffnen nun aber auch die konventionalistische Möglichkeit, falsifizierende Konsequenzen aus den für ATI-Annahmen negativen Evidenzen zu umgehen oder doch aufzuschieben, was für die Förderung des ATI-Erkenntnisfortschritts aber sowohl erwünschte (propulsiv-konstruktive) wie unerwünschte (degenerativ-immunisierende) Folgen haben kann. Die Schwierigkeit, (oberhalb bestimmter Minimalvoraussetzungen) aktuell wie prospektiv eindeutige empirische und außer-empirische Kriterien für die Beurteilung von Art, Richtung und Ausmaß dieses Fortschritts sowie seiner semantisch-pragmatischen Voraussetzungen positiv anzugeben (s. oben), behindert jedoch die Orientierung in künftigen Testentscheidungen. Schon zur Minimierung von Irrtumrisiken läßt dies deshalb die Propagierung und Anwendung eines mehrdimensional-pluralistischen Beurteilungssystems rational erscheinen, das auch mit methodologisch-epistemologischen Standards, Kriterien und Evaluationsmethoden nicht- oder prä-nomothetischer Art experimentiert. Die Verwendung weit heterogener Methodologieinhalte und -muster entspricht einmal dem aktuellen Problemstand (nicht nur) in diesem Bereich, der zuverlässige Prognosen über die Optimalität und Rationalität unterschied-

licher Formen der Forschungspraxis nicht zuläßt. Sie ermöglicht zudem erst jenes Maximum an epistemologisch und methodologisch relevanter Kritik, wie sie für die Steigerung der Leistungsfähigkeit alternativer Strategien und eine multidimensionale Beurteilung ihrer Forschungsleistungen erwünscht und erforderlich ist. Der weitere Verlauf des ATI-Programms wird damit nicht nur von der Bereitstellung eines möglichst großen Angebots an erklärungskräftigen oder doch entwicklungsfähigen Theorien des (schulischen) Lehrens und Lernens und der Organisierung und Regelung einer möglichst scharfen Konkurrenz zwischen ihnen abhängig gemacht, die derzeit besonders prekäre Theorienbildung in diesem Bereich bedarf zu ihrer Initiierung und Förderung vielmehr eines möglichst differenzierten und flexibel anwendbaren Ensembles von methodologischen Kriterien, Standards und Bewertungsverfahren, Entscheidungsprozeduren und Prüfmodellen, die Forschungsaktivitäten in weit unterschiedlichen Lehr-Lern-Situationen erst ermöglichen und nicht frühzeitig einschränken sollen. Die Problematisierung von Konstanz- und Homogenitätsannahmen, die im Verlauf des ATI-Programms bereits für übergeneralisierende Erklärungen und Vorhersagen individueller Lernleistungen erfolgte, erfaßt folglich auch den Universalitätsanspruch einer einheitlichen Standardmethodologie. An deren Stelle sollte vielmehr eine bereichsspezifisch variiende Beurteilung treten, welchen Methodologiesegmenten in welchen Problemlagen, bei welchem Theorieangebot und bei welchem theoretischen Entwicklungsstand welche den Erkenntnisfortschritt hemmenden oder fördernden Effekte zugesprochen werden können. Diese Konsequenz wird im übrigen auch in neueren Pluralismustheorien des Erkenntnisfortschritts gezogen, in denen nicht etwa das Auftreten, sondern das Ausbleiben divergierender bis al-

ternativer Theorie- und Methodologieprogramme als Krisensymptom von Wissenschaft gilt, das ein Stagnieren des Erkenntnisfortschritts signalisiert (vgl. zusammenfassend SPINNER 1974: 190ff.). Systematische und historische Problemanalysen der Wissenschaftsentwicklung in verschiedenen Gegenstandsbereichen und Disziplinen bestätigen diese Einschätzung zudem weitaus besser als die Analyseschemata monothetischer oder -methodologischer Fortschrittspostulate konkurrierender Beschreibungs- und Erklärungsversuche des Erkenntnisfortschritts. Rekonstruktion des bisherigen und die Normierung des künftigen Verlaufs des ATI-Programms legen denn auch die Übernahme dieser Sichtweise nahe.

### (3) Zur Anwendung von ATI-Annahmen

Die Anwendungsrelevanz der bisherigen ATI-Ergebnisse muß - nachdem größere Interaktionseffekte (noch) nicht gesichert werden konnten, eher gering eingeschätzt werden. Der im Bildungssystem aktivierte Verwertungsbedarf nach und Nutzungsanspruch auf Planungs-, Organisations- und Steuerungsdaten für schüleradaptive Instruktionsprozesse kann durch sie also zumindest mittelfristig nicht gedeckt werden. Dieser Bedarf leitet sich zwar aus der im ATI-Programm enthaltenen Problematisierung jener Konstanz-, Homogenitäts- und Stabilitätsannahmen ab, die den herkömmlichen Formen schulischen Lehrens und Lernens in Jahrgangsklassen eines vertikal gegliederten Bildungssystems zugrunde liegen und die inzwischen durch unterschiedlich weitreichende Differenzierungs- und Individualisierungsmaßnahmen abzufangen versucht wurde (vgl. ROEDER 1974). Die für deren effiziente Ausgestaltung und Handhabung erforderlichen diagnostischen Erhebungsinstrumente und technologisch-prognostischen Organisationsmittel fehlen indes und sind auch über eine nur heuristische Rezeption von ATI-Befunden nicht zu beschaffen (vgl. HOPF 1975). Substitu-

tiv können zwar andere Informationsquellen (wie z.B. Ergebnisse aus Modellversuchen zum zielerreichenden Lernen) genutzt werden, bedeuten aber einen Verzicht auf die Differenzierungsvorzüge der ATI-Perspektive oder sind an bestimmte Voraussetzungen (z.B. vergleichsweise große zeitliche Dispositionsfähigkeit oder den Nachweis von Hierarchiebeziehungen zwischen verschiedenen Vorkenntnisebenen) gebunden, die ihren Anwendungsnutzen mehr oder minder einschränken. Da mit wissenschaftlichen Auskünften über adaptiv-flexible Steuerungsmöglichkeiten im Unterricht, deren Voraussetzungen und Effekte anhand von ATI-Evidenzen also nicht zu rechnen ist und hinter den inzwischen erreichten Problemstand hinsichtlich der Reformbedürftigkeit eines starr durchgegliederten Bildungssystems auch nicht zurückgefallen werden kann, entsteht ein erhebliches Strukturierungs- und Orientierungsdefizit für die in ihm Handelnden bzw. die an seiner Innovation beteiligten Personen(gruppen). Diese lokalen Problemlösungslücken werden derzeit vermutlich nur durch den Plausibilitäts- und Evidenzcharakter subjektiver Didaktik-, Instruktions- und Schülertheorien überdeckt, deren Realitäts- und Rationalitätsgehalt nahezu unbekannt ist.

Für Praxisfelder mit besonders prekären Struktur-schwächen, wie sie sich beispielsweise bei der Einführung, Erprobung und Verbesserung binnendifferenzierender Lehr-Lern-Formen einstellen, ergeben sich aus dieser Einschätzung konstruktiv u.a. die folgenden Konsequenzen:

- Strategien der Ermöglichung zielerreichenden Lernens können über das von SALOMON (1972) beschriebene remediale ATI-Modell für die Differenzierung/Individualisierung im heterogenen Unterricht genutzt werden. Differenzierungsrelevant sind dabei die als Vorkenntnisse des jeweiligen Lernkriteriums (Unterrichtsziels)

Die Rekonstruktion der Zuordnungsregeln für die Elemente "Kriterium", "Unterrichtsmethode" und "Schülermerkmale" sowie der Verknüpfungsregeln zwischen diesen Elementen kann dann wie folgt genutzt werden:

- heuristisch für die wissenschaftsinterne Hypothesenbildung;
- evaluativ für den empirischen Vergleich des relativen Erklärungs- oder Vorhersagewertes analog rekonstruierbarer subjektiv-naiver und objektiv-wissenschaftlicher ATI-Annahmen, deren Bezug zum gleichen Objektbereich durch konsensual festgelegte Zuordnungsregeln gesichert werden kann (vgl. KLEITER 1974);
- explanativ für die Aufnahme subjektiver ATI-Annahmen in die Wenn-Komponente von Objektiv-Erklärungen des Individualisierungsverhaltens von Lehrern (vgl. HOPER et al. 1974) sowie des darüber vermittelten differentiellen Lernfortschritts von Schülern (vgl. WEINERT & GROEBEN 1975);
- korrektiv für die Präzisierung und Verbesserung der jeweils unterlegenen objektiven vs. subjektiven ATI-Annahme und die Optimierung ihres Anwendungsnutzens.

GROEBEN (1976) weist beispielsweise daraufhin, daß es bei der als weitgehend konstant nachgewiesenen Pünf-Dimensionalität impliziter Persönlichkeitstheorien von Lehrern unsinnig wäre, diese Komplexitätsschranke in dem auf die Schülerwahrnehmung bezogenen Kognitionssystem von Lehrern erheblich auszuweiten: ein Austausch zwischen subjektiven und wissenschaftlichen Schülerleistungserklärungen wäre bei der Vielzahl der im Forschungskontext bereits jetzt herangezogenen Erklärungsdimensionen gar nicht mehr möglich. Es erscheint denn auch unter einer anwendungsorientierten Forschungs- und Austauschperspektive durchaus plausibel, Beschreibung, Vorhersage und Erklärung von Schulleistungen auf einem der subjektiven Theorienbildung überhaupt noch kommensurablen Komplexitätsniveau vorzunehmen, wobei die Elastizität und Expansionsfähigkeit dieser subjektiv-vocativen Komplexitätsgrenzen durchaus erprobt werden kann. Kontrast, Konfrontation und Kritik subjektiver wie objektiver ATI-Aussagen können für beide Theorietypen deren Realitäts- und Rationalitätsgehalt erhöhen, wobei die Überlegenheit eines dieser Theorietypen (d.h. in der Regel des wissenschaftlichen) nicht dogmatisch vorausgesetzt wird, sondern als selbst noch überprüfens- und erklärenswert gilt.

Gelingt es, vor allem die subjektiven Voraussetzungen (wie z.B. kognitive Flexibilität, Strukturiertheit und Differenziertheit) für die Entwicklung und Selbststeuerung von Kognitionsinhalten (subjektive Theorien) zu erfassen, die eine adäquate Zuordnung der für unterschiedliche Schülergruppen jeweils optimalen Lernbedingungen bei gegebenen Unterrichtszielen beinhalten, so wäre damit auch die Generierung von objektiven ATI-Annahmen mit höherer Bewährungswahrscheinlichkeit zu erzielen als bei bisherigen Verknüpfungsversuchen einzelner Schülermerkmale und Unterrichtsvarianten. Hinzu kommt, daß die Erklärung des (instruktionsrelevanten) Individualisierungsverhaltens von Lehrern sowie des davon abhängigen Schülerlernfortschritts durch die Kenntnis dieser subjektiven Erklärungen verkürzt und erleichtert wird. Diese Vorteile wurden bislang allenfalls unsystematisch genutzt; ob und unter welchen Bedingungen sie gegeben sind und genutzt werden können, soll deshalb in einer Untersuchung zur Erklärung des Unterrichtserfolgs von Lehrern analysiert werden (vgl. TREIBER et al. 1976).

## 4. SPRACHENTWICKLUNG

---

### Gliederung

1. Der Ansatz Bernsteins
  - 1.1 Sprachliche Form der Intentionalität
  - 1.2 Sprachliche Form der Intentionalität
  - 1.3 Bernsteins Methodik
    - 1.3.1 Rezeption der Theorie Bernsteins
    - 1.3.2 Kritik an Bernstein
      - 1.3.2.1 Defizitmodell
      - 1.3.2.2 Sprechsituation als zusätzliche Determinante
      - 1.3.2.3 Charakteristiken der Sprechsituation
      - 1.3.2.4 Differenzmodell
  - 1.4 Neuformulierung der Theorie Bernsteins: *code vs. Sprachgebrauch*
    - 1.4.1 Definition
    - 1.4.2 Elaboriert/restringiert als Endpunkte einer kontinuierlichen Skala
    - 1.4.3 In welcher Situation wird am ehesten eine elaborierte Sprachvariante produziert?
  - 1.5 Sozialisationsinflüsse
    - 1.5.1 Entwicklung subjektiver Wahrscheinlichkeiten
    - 1.5.2 Interpretation der Situation
    - 1.5.3 Aspekte der Sprache
2. Der Ansatz Oevermanns
  - 2.1 Elementaristische Theorievariante
  - 2.2 Komplexe Version
  - 2.3. Kombination der beiden Aspekte bei Oevermann
  - 2.4 Kompensatorische Sprachprogramme
    - 2.4.1 Bohrreiter/Engelmann
    - 2.4.2 Cahagan
3. Das Verhältnis von Denken - Sprechen
  - 3.1 Denken als inneres Sprechen
  - 3.2 Unabhängigkeit von Denken und Sprechen
  - 3.3 Untersuchungen an sprachlich retardierten Kindern