

## DIAGNOSE DER SACHBEZOGENEN MOTIVATION IM SCHULUNTERRICHT

von Manfred Lehrke und Wilhelm F. Kempf

### 1. Problemstellung

Es wird kaum jemand der Behauptung widersprechen, "Motivation" habe eine entscheidende Bedeutung für das Lernen in der Schule. Kontrastierend zu dieser Einschätzung ist der relativ geringe Anteil an Überlegungen und empirischen Arbeiten zu Fragen von Motivation und Motivierung im pädagogischen Bereich. Die Beiträge der Motivationsforschung für die Verbesserung von Unterricht beschränken sich meist auf recht allgemeine Hinweise und Ratschläge. So lehren uns etwa die Erkenntnisse über Leistungsmotivation, daß Lernaufgaben einen mittleren Grad der Erreichbarkeit in einer gegebenen Lernsituation haben sollen. Von solchen Hinweisen, so richtig und wichtig sie sein mögen, ist es aber ein weiter Weg, bis sie präzisiert und in handhabbare didaktische Anweisungen transformiert worden sind. Dazu werden empirische Arbeiten benötigt.

Worin sind die Ursachen für diese relativ geringen Auswirkungen zu sehen? Eine Ursache dürfte in der Komplexität, bisweilen auch Mehrdeutigkeit und Verschwommenheit der verwendeten Motivationsbegriffe wie "Leistungsmotivation", "Neugier-motivation", "Interesse" liegen. Dem Konzept der "Leistungsmotivation" beispielsweise sind die verschiedensten Verhaltensweisen zuzuordnen, wenn sie nur durch die Auseinandersetzung mit einem Gütemaßstab charakterisiert werden können. HECKHAUSEN (1965, S. 604) unterscheidet sachbezogene, personbezogene und sozialbezogene Gütemaßstäbe, die konkreter oder abstrakter und mehr oder weniger differenziert sein können. Bei derartige Komplexität verwundert es nicht, wenn verschiedene Instrumente, die alle die Leistungsmotivation zu erfassen vorgeben, zum Teil völlig Verschiedenes messen (vergleiche MITCHELL 1961, HECKHAUSEN 1965, S. 605 ff.). Offensichtlich gibt es das allgemeine Leistungsmotiv nicht (vergleiche VEROFF 1973). Für den Bereich "Neugiermotivation" läßt sich Entsprechendes konstatieren (PEARSON 1970). "So sind gegenwärtig alle Bedingungen gegeben, es bei programmatischen Äußerungen bewenden zu lassen. Jeder hält Motivation für ungeheuer wichtig, aber kaum jemand weiß, was das näherhin sei, wie man

es greifen und was man konkret damit anfangen könnte." (HECK-HAUSEN 1973, S. 35).

Eine weitere Ursache ist in dem oft nicht ausreichenden Reflexionsgrad bezüglich der zu behandelnden erziehungswissenschaftlichen Fragestellungen zu suchen. Nicht selten bedient man sich in empirischen erziehungswissenschaftlichen Studien solcher Meßinstrumente, die man irgendwo vorfindet, die aber nicht speziell für diese Zwecke konstruiert worden sind.

Diese Instrumente zielen in der Regel auf eine "relativ überdauernde", "relativ konstante und generalisierte" Motivation, gleichsam auf eine Motivationsbereitschaft als Quasi-Persönlichkeitsmerkmal. Davon unterschieden wird die "aktualisierte Motivation" oder "Motiviertheit". Ob sich ein Schüler in einer bestimmten Lernsituation motiviert verhält, hängt außer von der "überdauernden" Motivation des betreffenden Schülers von Merkmalen der Situation ab. Dieser situative Einfluß kann offenbar viel größer sein, als man oft angenommen hat. Daraus ergibt sich als eine zentrale allgemeine Fragestellung für die Unterrichtsforschung, welche Merkmale Unterrichtssituationen (Lernsituationen) aufweisen müssen, damit sie günstige motivationale Prozesse und damit auch günstige Lernprozesse ermöglichen.

Es wird davon ausgegangen, daß Lernen möglichst von der Sache, vom Lernstoff her, motiviert sein sollte und nicht durch Druck, Zwang oder Belohnung von seiten des Lehrers oder der Eltern. Institutionalisiertes Lernen unterscheidet sich von außerschulischer, "natürlich" motivierter Informationsaufnahme und -verarbeitung dadurch, daß hier Lernen gefordert wird, und zwar mehr oder weniger unabhängig von den Bedürfnissen und Wünschen der Individuen. Ist beim Schüler aber keine sachbezogene, auf die Bearbeitung des Unterrichtsstoffes gerichtete Motivation vorhanden, so wird sein Handeln in der Regel von der Antizipation negativer Sanktionen geleitet sein. Er ist nicht motiviert, sich mit dem Stoff zu beschäftigen, sondern ist motiviert, jenen Sanktionen zu entgehen. Der Unterrichtsstoff ist dann ausschließlich als Mittel zu diesem Zweck von Bedeutung.

Derartige durch Abwesenheit von sachbezogener Motivation gekennzeichnete Unterrichtssituationen müssen eine Reihe negativer Konsequenzen haben. Diese Feststellung betrifft sowohl Ergebnisse kognitiven Lernens als auch eine Reihe übergeordneter

Bildungsziele. Wenn Lernaktivitäten nur durch äußeren Druck zustandekommen und aufrechterhalten werden, also in diesem Sinne situationsabhängig sind, kann beispielsweise ein "Lernen des Lernens" nicht erwartet werden. Ebenso wenig ist mit der Entwicklung von über den Unterricht oder die Schulzeit hinausreichenden Interessen oder wenigstens von Aufgeschlossenheit gegenüber den in der Schule behandelten Bereichen zu rechnen.

"Sachbezogene Motivation" soll die Zuwendung zu einem Lernstoff sein, die Bereitschaft, sich auch freiwillig mit diesem Sachverhalt zu beschäftigen, weitere Informationen dazu aufzunehmen und zu verarbeiten. Die Bezeichnung entstammt der Gegenüberstellung dieser und sachfremder oder nichtsachbezogener (also auf jeden Fall extrinsischer) Motivation beim Lernen. Der mit "sachbezogene Motivation" gemeinte Begriff unterscheidet sich von Begriffen wie "Leistungsmotivation" und "intrinsische Motivation" dadurch, daß jene sich nicht darauf beschränken, beobachtete bzw. beobachtbare Verhaltensweisen zu klassifizieren, sondern die Genese dieser Verhaltensweisen zum Klassifikationsmerkmal machen. So muß ein Verhalten, das als ein leistungsmotiviertes gelten soll, durch die Auseinandersetzung mit einem Gütemaßstab bewirkt worden sein. Intrinsische Motivation liegt nur dann vor, wenn ein Verhalten nicht wegen möglicher Folgen gezeigt wird, sondern weil die Aktivität als angenehm empfunden wird. Sachbezogene Motivation ist also nicht identisch mit intrinsischer Motivation, sondern ist allgemeiner. Abgesehen von der Schwierigkeit, intrinsische Motivation "rein" zu erfassen, erscheint es weder realistisch noch wünschenswert, in der Schule nur intrinsisch, also durch Lernaktivitäten selbst und nicht mittels ihrer Folgen, motivieren zu wollen. Extrinsische Motivationsquellen unterscheiden sich nach Art, Stärke und nach ihrer Bedeutung für das Lernen erheblich, sie sind keineswegs stets lernhemmend. Die mögliche Mitbedingtheit durch extrinsische Faktoren ist ein wesentlicher Aspekt sachbezogenen motivierten Lernens. Einer als sachmotiviert geltenden Verhaltensweise läßt sich nicht von vornherein ansehen, wodurch sie angeregt worden ist. Weiter braucht nicht davon ausgegangen zu werden, daß das gleiche Verhalten bei verschiedenen untersuchten Schülern auf die gleiche Weise zustandekommt. Durch welche Art der Motivierung eine Änderung des Zuwendungsverhaltens zu einem Lernstoff bei Schülern erreicht worden ist, wird zu einer Frage des experimentellen Designs. Dadurch können ganz unterschiedliche Motivierungsansätze erprobt und ihre Effekte miteinander verglichen werden.

Für Untersuchungen in diesem Bereich wird ein Meßinstrument benötigt, daß bei der Unterrichtsforschung, Curriculumentwicklung und -evaluation, aber auch von Lehrern leicht handhabbar ist. Es darf nur so wenig aufwendig sein, daß es den Unterricht nicht behindert und bei Bedarf in derselben Klasse öfter eingesetzt werden kann. Außerdem soll es für verschiedene Schulstufen, Fächer, Inhalte, Methoden und Lehrer verwendbar sein und vergleichbare Werte liefern. Eine weitere Forderung ist die nach individuellen Meßwerten, damit die Auswirkungen technologischer Eingriffe in den Unterricht auf den einzelnen Schüler erfaßbar sind. Gemäß den skizzierten Fragestellungen soll nicht die überdauernde, sondern die aktualisierte sachbezogene Motivation erfaßt werden.

Mit der oben gegebenen Definition der sachbezogenen Motivation im Unterricht als der Zuwendung zu den Lernstoffen wird eine Vielzahl von Verhaltensweisen unter dem Oberbegriff der "sachbezogenen Motivation" zusammengefaßt und als Indiz für sachbezogenen Motiviertheit betrachtet. Wer im Unterricht Interesse an den Lerninhalten zeigt, wer Spaß daran findet, kann ebenso als sachbezogen motiviert gelten, wie jener Schüler, der sich auch außerhalb des Unterrichts und freiwillig mit den Lernstoffen weiterbeschäftigt.

Der Zusammenfassung so unterschiedlicher Verhaltensweisen unter einem gemeinsamen Begriff der sachbezogenen Motivation liegt dabei implizit die Annahme zugrunde, daß man die Schüler nach dem Ausmaß ordnen kann, in welchem sie sachbezogen motiviert sind, und daß von je zwei Schülern derjenige, der z. B. eher Spaß an den Unterrichtsinhalten findet, zugleich auch stärker dazu tendiert, sich freiwillig und außerhalb des Unterrichts damit weiterzubeschäftigen, usw. Die Problematik ist dieselbe, wie sie sich auch aus dem Gebrauch psychologischer Eigenschaftsbegriffe (vergleiche KEMPF 1974a, b) ergibt: die Einführung und Verwendung des Begriffs der sachbezogenen Motivation setzen das Vorliegen einer empirischen Gesetzmäßigkeit voraus. Diese Gesetzmäßigkeit besteht im wesentlichen in der Generalisierbarkeit der Aussagen, welche über die sachbezogene Motivation getroffen werden. Sie besteht darin, daß die Aussagen über die Motivation, welche aus einer Verhaltensweise erschlossen werden, auch auf andere Verhaltensweisen übertragbar sein müssen. Wird mehr als nur eine Verhaltensweise als Indiz für sachbezogene Motiviertheit verstanden, derart, daß von zwei Schülern jener als motivierter gilt, der diese Verhaltensweise mit größerer Wahrscheinlichkeit

zeigt, so müssen alle diese Verhaltenswahrscheinlichkeiten durch streng monoton wachsende Transformationen ineinander überführbar sein. Soll nicht nur eine Rangordnung der Schüler festgelegt werden, sondern strebt man eine leistungsfähigere Motivationskala an, so müssen auch noch mehrstellige Relationen zwischen den Schülern definiert werden, z. B. indem man bestimmte Annahmen über die funktionale Form des (streng monoton wachsenden) Zusammenhangs zwischen den obengenannten Verhaltenswahrscheinlichkeiten und der Motiviertheit der Schüler postuliert. Wie man heute zeigen kann, müssen diese Annahmen allerdings bestimmten wissenschaftstheoretischen und mathematisch-statistischen Voraussetzungen (spezifische Objektivität, bedingte Inferenz) genügen, wodurch die Menge der möglichen Funktionen auf eine kleine und überschaubare Anzahl reduziert wird (vergleiche KEMPF 1974b).

## 2. Versuchsanordnung

Mit der vorliegenden Studie wurde der Versuch unternommen, sachbezogene Motivation im Schulunterricht mit Hilfe von Fragebogen zu messen. Die darin verwendeten Fragen sind alle auf denselben, jeweils einem bestimmten Unterrichtsabschnitt zuzuordnenden Unterrichtsstoff beziehbar, ohne daß dieser Inhalt in den Items selbst erscheint. Der jeweilige Bezug wird durch die Instruktion hergestellt, womit die Anwendbarkeit der Fragebogen auf verschiedene Unterrichtsformen und -inhalte und die Vergleichbarkeit der daraus erhaltenen Ergebnisse ermöglicht wird.

Insgesamt wurden 22 Fragen zusammengestellt, die Verhaltensweisen wie "Interesse am Unterrichtsstoff zeigen", "sich freiwillig mit ihm beschäftigen" usw. betreffen. Die Items haben die Form von Aussagen, die für den einzelnen Schüler entweder zutreffen oder nicht zutreffen. Er wird aufgefordert, die für ihn jeweils zutreffende Antwort "stimmt" oder "stimmt nicht" anzukreuzen.

Die Hälfte der Items ist "positiv" formuliert, d. h. ihre Bejahung wird als Indiz für sachbezogene Motiviertheit angesehen. Es sind dies die Fragen Nr. 1, 2, 3, 4, 9, 10, 11, 13, 18, 21 und 22. Die andere Hälfte der Items ist "negativ" formuliert, d. h. ihre Verneinung wird als Indiz für sachbezogene Motiviertheit betrachtet. Es sind dies die Fragen Nr. 5, 6, 7, 8, 12, 14, 15, 16, 17, 19 und 20.

Der Wortlaut der Fragen ist:

- Frage 1: Es gab Fragen bzw. Themen, die mich besonders interessiert haben.
- Frage 2: Bestimmte Dinge sind meinem Interesse nach zu kurz behandelt worden.
- Frage 3: Ich habe auch außerhalb des Unterrichts über manche Dinge nachgedacht, die wir zuletzt gelernt haben.
- Frage 4: Ich habe mich hauptsächlich deshalb am Unterricht beteiligt, weil mir das Thema Spaß gemacht hat.
- Frage 5: Ich habe weder in Lexika noch woanders nachgeschlagen, um mehr Informationen über das behandelte Gebiet zu bekommen.
- Frage 6: Ich möchte mich in den nächsten Stunden lieber nicht mehr mit ähnlichen Fragen beschäftigen.
- Frage 7: Während des Unterrichts habe ich nicht gemerkt, daß es Spaß machen kann, so etwas zu lernen.
- Frage 8: Wenn ich jetzt etwas über die Dinge, die wir im Unterricht behandelt haben, in der Zeitung finden würde, würde ich es wahrscheinlich nicht lesen.
- Frage 9: Es gab keine Fragen, mit denen ich mich meinem Interesse nach zu lange beschäftigen mußte.
- Frage 10: Zu diesem Gebiet finde ich Hausaufgaben gut.
- Frage 11: Ich habe das Gefühl, für mich selbst etwas dazugelernt zu haben.
- Frage 12: Noch mehr möchte ich nicht über dieses Thema erfahren.
- Frage 13: Die Schule würde mir mehr Spaß machen, wenn wir öfter solche Dinge behandeln würden.
- Frage 14: Wenn mir jemand ein Buch über die zuletzt behandelten Themen schenken würde, würde ich mich darüber nicht besonders freuen.
- Frage 15: Ich hätte mich nie mit diesem Gebiet beschäftigt, wenn ich nicht gemußt hätte.
- Frage 16: Einiges hat mich ausgesprochen gelangweilt.
- Frage 17: Mein Interesse an Naturwissenschaften ist nicht größer geworden, seit wir diesen Stoff durchgenommen haben.
- Frage 18: Ich wünschte, es gäbe bald eine Fernsehsendung über dieses Thema.
- Frage 19: Ich habe mich nur deshalb am Unterricht beteiligt, weil ich keine schlechte Zensur bekommen möchte.

Frage 20: So interessant war es nie, daß ich es schade fand, als es klingelte und die Stunde vorbei war.

Frage 21: Ich habe mit Freunden, Eltern oder Geschwistern über Dinge aus diesem Gebiet gesprochen.

Frage 22: Mit welchen Themen hätte ich mich auch freiwillig gerne beschäftigt, z. B. in einer Arbeitsgruppe ohne Zensuren oder zu Hause.

Zur Überprüfung der Annahme, daß alle diese Fragen zur Messung derselben latenten Variable, nämlich der aktualisierten sachbezogenen Motivation der Schüler, geeignet sind, ist zu zeigen, daß die verschiedenen Fragen tatsächlich alle zu derselben Rangordnung der Schüler führen, d. h. daß die Aussagen, welche aus der Beantwortung eines Items (oder einer Teilgruppe von Items) über die sachbezogene Motiviertheit der Schüler erschlossen werden, nicht in Widerspruch zu den Aussagen stehen, welche aus einem anderen Item (oder einer anderen Itemgruppe) folgen, sondern sich von diesen allenfalls in ihrer Genauigkeit unterscheiden.

In der testtheoretischen Literatur wird diese Problematik unter dem Stichwort "Homogenität" abgehandelt. Sie kann immer nur relativ zu einem gegebenen Meßmodell untersucht werden, da unterschiedliche Meßmodelle die Homogenität verschieden definieren (vergleiche FISCHER 1968). Zwar könnte man auf die weitere Spezifikation unseres - bisher nur verbal formulierten - Meßmodells verzichten und die Rangordnung der Schüler aus der Anzahl der im Sinne sachbezogener Motiviertheit beantworteten Items schätzen. Wie KEMPF (1974b) zeigt, führt diese Vorgangsweise auch zu einer erwartungstreuen Schätzung der Rangordnung, doch ist sie mit einem unkontrollierbaren Informationsverlust verbunden, wenn diese Anzahl keine erschöpfende Statistik ist. Geht man davon aus, daß die Motiviertheit der Schüler durch die Beantwortung der Fragebogenitems unbeeinflusst bleibt, so können die Items als lokal unabhängig betrachtet werden. Die Forderung nach erschöpfenden Statistiken der Form

$$a_{vo} = \text{Anzahl der Fragebogenitems, welche der Schüler } v \text{ im Sinne sachbezogener Motiviertheit beantwortet}$$

führt dann notwendigerweise zu RASCHs (1960) logistischem Testmodell, das der vorliegenden Studie zugrunde gelegt wurde. Un-

ter den obengenannten Voraussetzungen ist das Testmodell von RASCH auch notwendig und hinreichend für die Gewährleistung spezifischer Objektivität und für die Anwendbarkeit bedingter Inferenzmethoden (vergleiche KEMPF 1974b).

Wenn Änderungen der Motiviertheit von Unterrichtsabschnitt zu Unterrichtsabschnitt gemessen werden sollen, so setzt ein solches Vorgehen voraus, daß die Items universell, d.h. auch bei inhaltlich und methodisch unterschiedlichem Unterricht anwendbar sind und vergleichbare Werte liefern, also stets dasselbe messen. Die einmal nachgewiesene Homogenität müßte auch dann erhalten bleiben, wenn die Items bei unterschiedlichen Bedingungen eingesetzt werden. In der vorliegenden Untersuchung wurden vier einander teilweise überlappende Einzeltests konstruiert, in denen Fragen doppelt, also in zwei Einzeltests zugleich, vorkamen.

Als Stichprobe standen zunächst 356 Schüler aus drei neunten und zwölf zehnten Klassen zur Verfügung. Drei von ihnen waren Hauptschulklassen, sieben aus Realschulen und fünf aus Gymnasien. Beim Durchgang mit dem vierten Einzeltest waren noch elf der Klassen mit 252 Schülern dabei.

Die Fragebogen wurden im Physikunterricht eingesetzt, die drei ersten während einer gut zwei Monate dauernden Unterrichtseinheit über den Modellbegriff, der vierte nach einem knapp zwei-monatigen Unterricht in Kleingruppen über Themen aus dem Bereich Elektronik. Die Teilthemen der Unterrichtseinheit über den Modellbegriff wurden in den verschiedenen Klassen in unterschiedlicher Reihenfolge und Ausführlichkeit dargeboten. Die Gruppen des anschließenden Unterrichts behandelten ganz unterschiedliche Fragen: Experimente mit Transistorschaltungen, logische Schaltungen, Spannungsverstärker, Kauf von Phono- und Fernsehgeräten.

Jeweils vor Beginn eines neuen Unterrichtsabschnitts war ein Fragebogen zu bearbeiten, der sich auf den vorhergehenden Unterricht bezog. Da alle Schüler die Einzeltests derselben Reihenfolge erhielten, stammen die zu analysierenden Daten jedes Einzeltests aus wenigstens nach Länge und Inhalt unterschiedlichem Unterricht, der von verschiedenen Lehrern gehalten wurde.

Die Zuordnung der Fragebogenitems zu den vier Tests und ihre Positionen innerhalb der Einzeltests sind aus Tab. 1 ersichtlich.



Frage Nr.	Item-Nr. in Einzeltest 1	Item-Nr. in Einzeltest 2	Item-Nr. in Einzeltest 3	Item-Nr. in Einzeltest 4
1	1	2	-	-
2	2	-	6	-
3	3	-	7	-
4	4	-	-	4
5	5	-	-	-
6	6	1	-	-
7	-	3	10	-
8	-	4	-	6
9	-	5	-	7
10	-	6	-	-
11	7	-	1	-
12	-	7	2	-
13	-	-	3	8
14	-	-	4	10
15	-	-	5	-
16	8	-	-	2
17	-	8	-	3
18	-	9	-	1
19	-	11	8	-
20	-	-	-	5
21	9	-	9	-
22	-	10	-	9

Tab. 1: Zuordnung der Items zu den Einzeltests

Die schriftliche Instruktion für den ersten Einzeltest lautete:

- (1) Wir bitten Dich um Deine Mitarbeit bei der Feststellung von Schülerinteressen. Dieser Fragebogen hat nichts mit Zensuren zu tun, und der Lehrer erfährt nicht, wer was angekreuzt hat.
- (2) Der Fragebogen enthält einige Aussagen über Interessen, die für Dich entweder zutreffen oder nicht zutreffen. Die Aufgabe besteht darin, die Aussagen durchzulesen und jedesmal entweder "stimmt" oder "stimmt nicht" anzukreuzen. Arbeite dabei bitte allein und lasse keine Frage aus, weil man nur dann genügend über die verschiedenen Interessen erfährt.
- (3) Der Fragebogen bezieht sich auf folgenden Unterrichtsstoff: (wird vom Lehrer bekanntgegeben).

Das Thema des Unterrichts wurde an der vorgesehenen Stelle durch die Schüler eingesetzt. Die Instruktionen der drei anderen Einzeltests konnten durch Verweise auf die Durchführung der vorangegangenen etwas kürzer gehalten werden.

### 3. Ergebnisse

Eine wesentliche Modelleigenschaft des Testmodells von RASCH besteht in der spezifischen Objektivität der Aussagen über die Personen und/oder die Items. Werden z. B. die Item-Schwierigkeiten mit Hilfe der bedingten Maximum-Likelihood-Methode (CML-Methode) geschätzt, so sind die daraus resultierenden Schätzfunktionen von der Wahl der Personentichprobe unabhängig. Das bedeutet, daß man die Itemschwierigkeiten aus beliebigen Teilstichproben von Personen schätzen und die Schätzungen miteinander vergleichen kann. Ist das Modell erfüllt, d. h. messen alle Items dieselbe latente Variable, dann dürfen sich die Schätzungen nur unwesentlich voneinander und von der Schätzung aus der Gesamtstichprobe unterscheiden.

Eine Möglichkeit, diesen Vergleich durchzuführen, besteht in der graphischen Gegenüberstellung. Dabei werden die Schätzungen aus der Gesamtstichprobe auf der Abszisse eines Koordinatensystems aufgetragen, die Schätzungen aus den Teilstichproben auf der Ordinate. Bei Geltung des Modells müssen die durch die Koordinaten festgelegten Punkte annähernd auf der Geraden

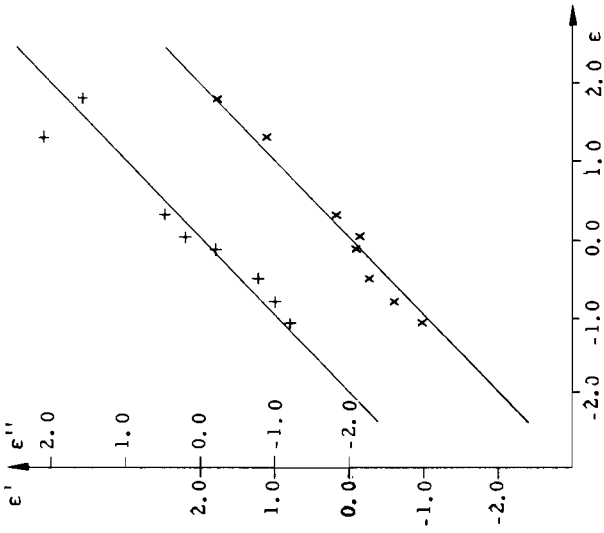


Abb. 1: Graphische Modellkontrolle Einzeltest 1, unselektiert

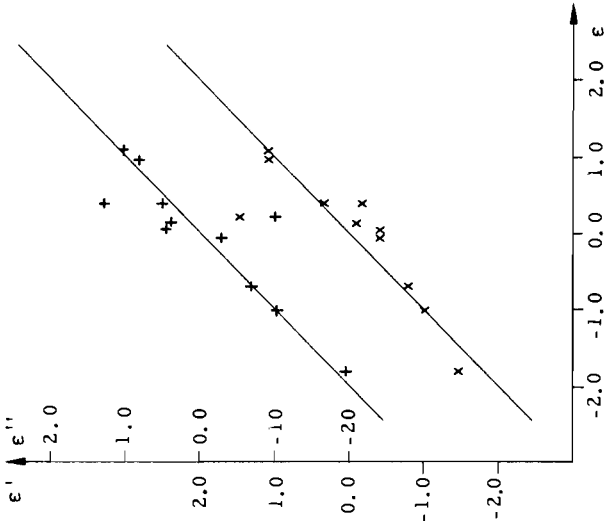


Abb. 2: Graphische Modellkontrolle Einzeltest 2, unselektiert

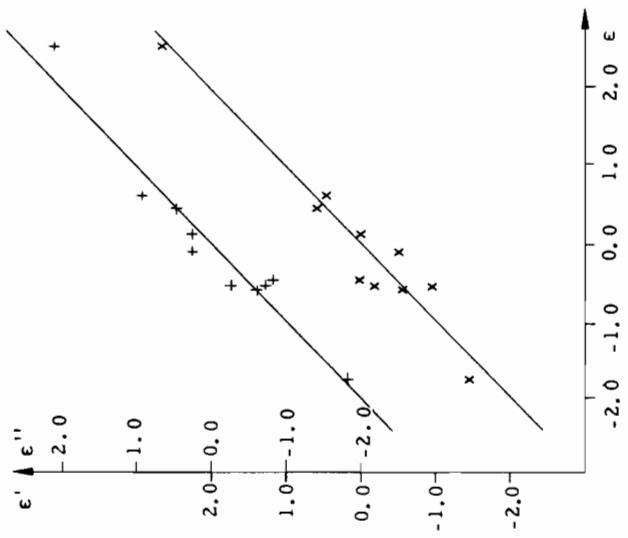


Abb. 3: Graphische Modellkontrolle Einzeltest 3, unselektiert

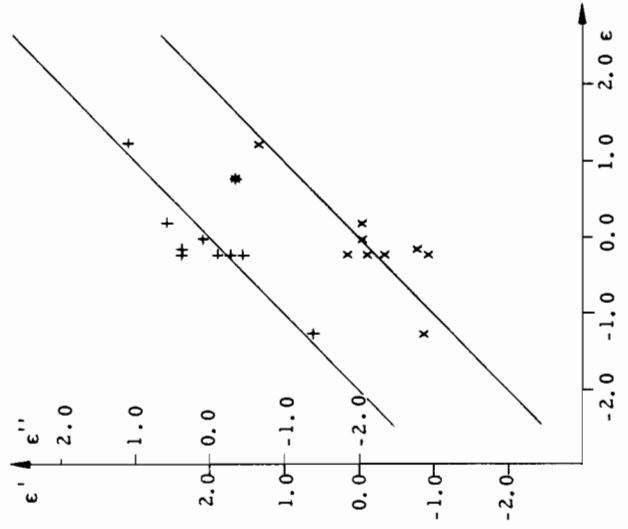


Abb. 4: Graphische Modellkontrolle Einzeltest 4, unselektiert

liegen, die mit einem Neigungswinkel von 45 Grad durch den Ursprung des Koordinatensystems führt.

Als statistisches Prüfungsverfahren kann der Vergleich in Form eines bedingten Likelihood-Quotienten-Tests (vergleiche ANDERSEN 1974) vorgenommen werden, der eine asymptotisch CHI-Quadrat verteilte Testgröße liefert und gegenüber der graphischen Modellkontrolle deswegen vorzuziehen ist, weil letztere keine Anhaltspunkte dafür bietet, welcher Grad der Abweichung eines Items noch zu tolerieren ist. In der vorliegenden Arbeit wird die graphische Modellkontrolle daher nur zur Illustration herangezogen.

Die Zuordnung der Schüler zu den Teilstichproben erfolgte auf der Grundlage ihrer Rohscores  $a_{VO}$ , d. h. es wurden stark motivierte und weniger stark motivierte Schüler einander gegenübergestellt. Diese Form des Modelltests wird der formalen Struktur des RASCH-Modells am besten gerecht.

Die Ergebnisse der ersten Itemanalyse, in der jeweils alle Items eines Einzeltests gemeinsam untersucht wurden, sind in Abb. 1 bis 4 dargestellt. Die zugehörigen Resultate der Modelltests sind in Tab. 2 zusammengefaßt.

Die Modellanpassung ist durchweg schlecht. Der verwendete Itempool ist in seiner vorliegenden Form nicht dazu geeignet, die sachbezogene Motiviertheit im Schulunterricht zu definieren. Zumindest einzelne Fragen sprechen über die sachbezogene Motiviertheit hinaus auch noch andere Reaktionstendenzen der Schüler an. So könnte die Ursache der mangelnden Modellanpassung z. B. in der Wirksamkeit von Antworttendenzen (z. B. Aquieszenz) zu suchen sein; denn alle vier Einzeltests enthielten sowohl positiv als auch negativ formulierte Fragen. Der störende Einfluß, welchen solche Antworttendenzen auf die Homogenität der Items ausüben, müßte jedoch verschwinden, wenn positiv und negativ formulierte Items getrennt voneinander analysiert werden. Wie die in Tab. 3 zusammengefaßten Modelltest-Ergebnisse zeigen, führt aber auch die Trennung von positiv und negativ formulierten Fragen nur zu einer unwesentlichen Verbesserung der Modellanpassung. Immer noch sind 5 von 8 Modelltests signifikant, einer liegt knapp an der Grenze der Signifikanz, und nur bei den positiv formulierten Items des dritten Einzeltests sowie bei den negativ formulierten Items des vierten Einzeltests zeigt sich eine gute Modellanpassung.

Einzeltest	Fragen Nr.	Modelltest	df	p
1	1, 2, 3, 4, 5, 6, 11, 16, 21.	19.544	8	$p > .98$
2	1, 6, 7, 8, 9, 10, 12, 17, 19, 22.	126.024	10	$p > .999$
3	2, 3, 7, 11, 12, 13, 14, 15, 19, 21.	16.430	9	$.90 < p < .95$
4	4, 8, 9, 13, 14, 16, 17, 18, 20, 22.	81.426	9	$p > .999$

Tab. 2: Modelltestergebnisse der unselegierten Einzeltests

Positiv formulierte Items				
Einzeltest	Fragen Nr.	Modelltest	df	p
1	1, 2, 3, 4, 11, 21	9.732	5	$.90 < p < .95$
2	1, 9, 10, 18, 22	42.244	4	$p > .999$
3	2, 3, <sup>+</sup> <del>11</del> , 13, 21	1.674	3	$.30 < p < .50$
4	4, 9, 13, 18, 22	68.005	4	$p > .999$
Negativ formulierte Items				
Einzeltest	Fragen Nr.	Modelltest	df	p
1	5, 6, 16	8.886	2	$p > .98$
2	6, 7, 8, 12, 17, 13	26.014	5	$p > .999$
3	7, 12, 14, 15, 19	10.368	4	$p > .95$
4	8, 14, 16, 17, 20	3.036	4	$.30 < p < .50$

Tab. 3: Modelltestergebnisse der unselegierten Einzeltests, getrennt für positiv und negativ formulierte Items

+) Item 11 (Nr. 1 in Einzeltest 3) zu leicht

Zusammenfassend muß daher festgestellt werden, daß über die eventuelle Wirksamkeit von Antworttendenzen hinaus auch noch andere Störfaktoren im Spiel sind. Auf dem Weg einer Inhaltsanalyse der Items kann der Versuch unternommen werden, solche Störfaktoren aufzudecken, die entsprechenden Items aus den Einzeltests auszuschneiden und in neuerlichen Itemanalysen zu untersuchen, ob die verbleibenden Items einen homogenen Itemsatz bilden, d. h. ob die in ihnen angesprochenen Verhaltensweisen als Indiz für eine Variable der sachbezogenen Motiviertheit dienen können.

Soll das Konzept der sachbezogenen Motiviertheit dabei keiner inhaltlichen Einschränkung unterliegen, so wird man zunächst versuchen müssen, solche Items auszuschneiden, an welchen sich Formulierungsmängel feststellen lassen. Solche Items sind die Fragen 10 und 19, die unterschiedlich interpretiert werden können.

- (a) Die Feststellung "Zu diesem Thema finde ich Hausaufgaben gut" (Frage 10) kann entweder bedeuten, daß der Schüler lieber eine Hausaufgabe aus diesem Thema schreiben möchte als auch einem anderen, oder aber, daß der Schüler die Hausaufgabe gegenüber der Alternative "keine Hausaufgabe" bevorzugt. Je nachdem, wie der Schüler die Frage interpretiert, wird ihm die Zustimmung zu der Feststellung leichtfallen oder nicht. Sind zwei Schüler gleich stark motiviert, so wird dennoch jener mit größerer Wahrscheinlichkeit zustimmen, der die Frage im erstgenannten Sinn interpretiert hat.
- (b) Stimmt ein Schüler der Feststellung "Ich habe mich nur deshalb am Unterricht beteiligt, weil ich keine schlechte Zensur bekommen möchte" (Frage 19) nicht zu, so muß dies kein Hinweis auf sachbezogene Motivation sein. Es kann auch bedeuten, daß die Vermeidung schlechter Zensuren für den Schüler kein Motiv für Beteiligung am Unterricht darstellt oder daß dafür zwar andere, aber ebenfalls nicht-sachbezogene Motive ausschlaggebend waren.

In den nachfolgenden Itemanalysen, in denen weiter positiv und negativ formulierte Items getrennt voneinander untersucht wurden, mußten auch noch ausgeschieden werden:

- aus Einzeltest 1 die Fragen 2 und 5,
- aus Einzeltest 2 die Fragen 8 und 9,

- aus Einzeltest 3 die Frage 14,
- aus Einzeltest 4 die Frage 9.

Tab. 4 zeigt, daß die danach verbliebenen Items homogene Subgruppen bilden. Bei Einzeltest 1 war für die negativ formulierten Items 6 und 16 eine Homogenitätsprüfung - für die mindestens 3 Items vorhanden sein müssen - nicht möglich. Die Ergebnisse für die Positiv formulierten Items aus Einzeltest 3 sowie für die negativ formulierten Items aus Einzeltest 4 stehen bereits in Tab. 3, weil hier keine Items mehr selektiert werden mußten. Die graphische Darstellung der Modelltests findet sich in Abb. 5 bis Abb. 8.

Während die Fragen 10 und 19 wegen ihrer mißverständlichen Formulierungen ungeeignet sind, sprechen die restlichen ausgeschiedenen Fragen solche Verhaltensweisen an, die nur unter bestimmten Umständen als Indiz für sachbezogene Motiviertheit herangezogen werden können. Bei zwei Items zeigte sich eine

Positiv formulierte Items				
Subtest Nr.	Fragen Nr.	Modelltest	df	p
1	1, 3, 4, 11, 21	3.348	4	.30 < p < .50
2	1, 18, 22	1.729	2	.50 < p < .70
3	2, 3, <sup>11</sup> 11, 13, 21	1.674	3	.30 < p < .50
4	4, 13, 18, 22	3.418	3	0.50 < p < 0.70
Negativ formulierte Items				
Subtest Nr.	Fragen Nr.	Modelltest	df	p
1	6, 16	—	—	—
2	6, 7, 12, 17	4.489	3	.75 < p < .80
3	7, 12, 15	1.443	2	.50 < p < .70
4	8, 14, 16, 17, 20	3.036	4	.30 < p < .50

Tab. 4: Modellergebnisse nach der Itemselektion, getrennt für positiv und negativ formulierte Items

+) Item 11 (Nr. 1 in Einzeltest 3) zu leicht



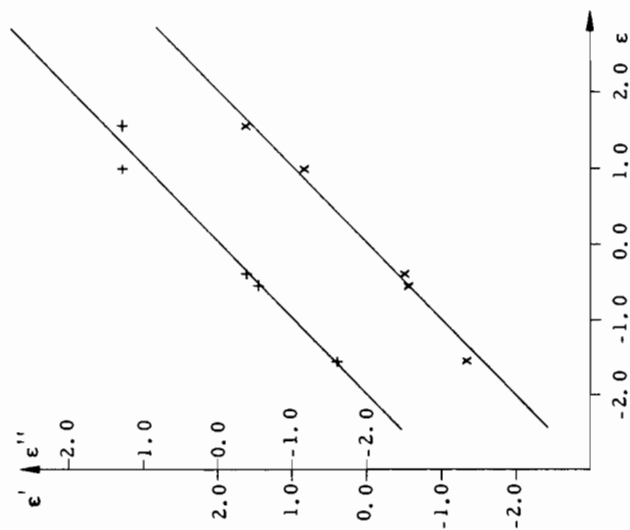


Abb. 5: Graphische Modellkontrolle Einzeltest 1, selektiert, positiv formulierte Items

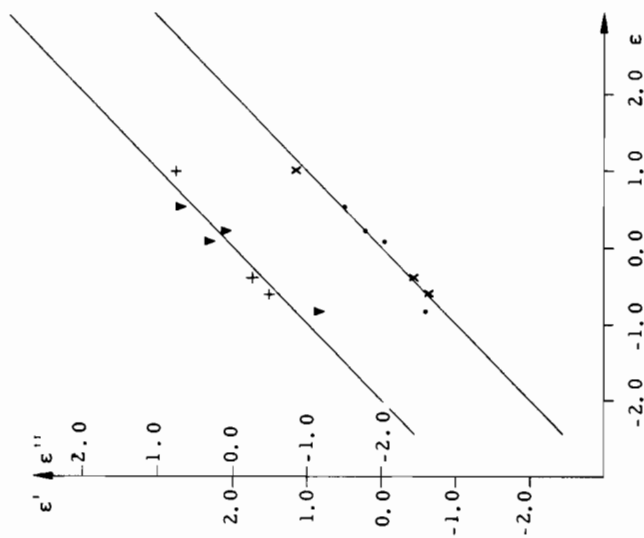


Abb. 6: Graphische Modellkontrolle Einzeltest 2, selektiert, getrennt für positiv und negativ formulierte Items

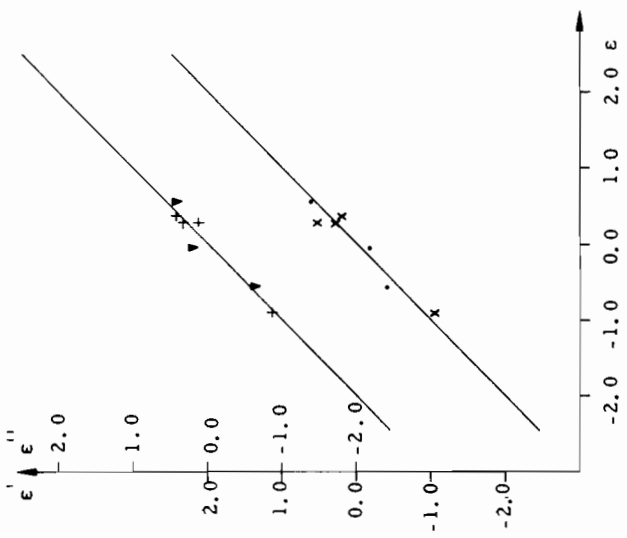


Abb. 7: Graphische Modellkontrollen Einzeltest 3, selektiert, getrennt für positiv und negativ formulierte Items

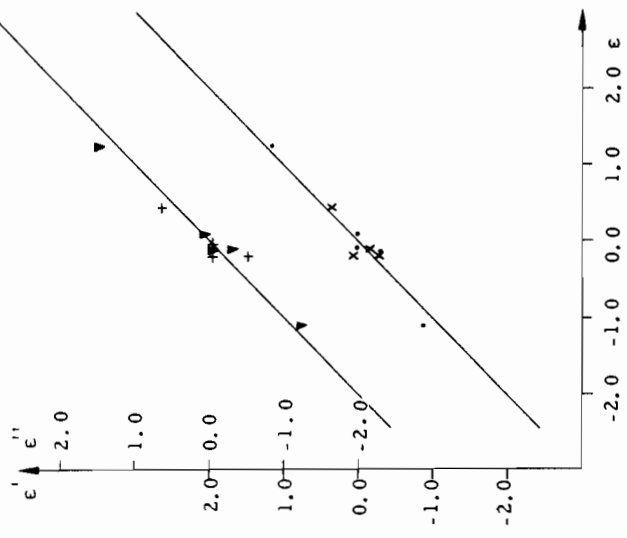


Abb. 8: Graphische Modellkontrollen Einzeltest 4, selektiert, getrennt für positiv und negativ formulierte Items

Abhängigkeit ihrer Schwierigkeiten von der Variabilität des Unterrichts, bei drei Items würden die Antworten mitbestimmt durch unterschiedliche Voraussetzungen bei den Schülern. Auf diese beiden Fälle soll näher eingegangen werden, weil daraus bestimmte Folgerungen gezogen werden können.

Die Fragen 2 und 9 wurden ausgeschieden, weil sie nicht nur die Motiviertheit an sich ansprechen, sondern eine ausgewogene Relation zwischen der Motiviertheit einerseits und dem im Unterricht realisierten Informationsangebot andererseits. Je stärker der Unterricht in Dauer und Ausführlichkeit der Behandlung der Unterrichtsinhalte variiert, desto weniger sind die Items zur Messung der sachbezogenen Motivation geeignet.

Wie bereits oben ausgeführt, wurden die Teilthemen über den Modellbegriff in den verschiedenen Schulklassen unterschiedlich ausführlich behandelt. Diese Variation des Unterrichts war bei jenen Themen, auf welche sich der erste Einzeltest bezog, sehr stark und wurde durch die Variabilität der Behandlung der Themen, auf welche sich der zweite Test bezog, wieder weitgehend ausgeglichen, so daß der Unterricht zu jenen Themen, auf die sich der dritte Test bezog, relativ homogen gestaltet war. Tatsächlich mußte die Frage 2 nur im 1. Test ausgeschieden werden, während sie im 3. Test beibehalten werden konnte. Frage 9 mußte sowohl im 2. Test als auch im 4. Test ausgeschieden werden. Der 4. Test hatte sich auf einen Gruppenunterricht bezogen, dessen Variabilität über die verschiedenen Gruppen ebenfalls recht stark war.

Die Fragen 5, 8 und 14 sprechen neben der Motivation gewisse Randbedingungen des Verhaltens an, die vereinfachend als "Lesegewohnheiten" bezeichnet werden sollen. Dabei fällt auf, daß die Fragen 8 und 14 nur aus dem 2. bzw. 3. Einzeltest ausgeschieden werden mußten, nicht jedoch aus dem 4. Test. Die spezielle Form des Gruppenunterrichts führte offensichtlich zu einer Stabilisierung der Randbedingungen, so daß diese sich nicht mehr als Störfaktoren auswirken konnten. Bei diesem Unterricht wurden Gruppen von je 4 Schülern gebildet, die die Themen weitgehend selbständig bearbeiteten. Zu den Arbeitsunterlagen, welche den Schülern zur Verfügung gestellt wurden, gehörte unter anderem eine Textsammlung, von der ausgiebig Gebrauch zu machen war. Offensichtlich hat erst diese aktive Form der Informationsaufnahme durch Lesen einen Transfer der Motiviertheit auf solche Verhaltensweisen wie Bücher- und Zeitungslesen

Einzeltest	Fragen Nr.	Modelltest	df	p
1	1, 3, 4, 6, 11, 16, 21	21.169	6	$p > .995$
2	1, 6, 7, 12, 17, 18, 22	7.479	6	$.7 < p < .75$
3	<sup>+) </sup> 2, 3, 7, <del>11</del> , 12, 13, 15, 21	13.733	6	$p > .95$
4	4, 8, 13, 14, 16, 17, 18, 20, 22	30.981	8	$p > .999$

Tab. 5: Modelltestergebnisse nach der Itemselektion für positiv und negativ formulierte Items gemeinsam

+) Item 11 (Nr. 1 in Einzeltest 3) zu leicht

bewirkt, wie sie in den Fragen 8 und 14 angesprochen werden. Durch die Arbeit mit der Textsammlung kam es zur Ausbildung eines speziellen Zuwendungsverhaltens zu Texten über den Unterrichtsstoff, dessen Stärke durch die sachbezogene Motiviertheit des Schülers bestimmt ist, während die allgemeinen Lesegewohnheiten des Schülers ihren Einfluß eingebüßt haben.

Nach der Auffindung in sich homogener Gruppen von positiv und negativ formulierten Items kann nun erneut die Frage nach der eventuellen Wirksamkeit von Antworttendenzen gestellt werden. Dazu ist zu überprüfen, ob die in den Einzeltests verbliebenen positiv formulierten und negativ formulierten Items auch zusammengenommen homogene Itemsätze bilden. Die Ergebnisse dieser Itemanalysen sind in Tab. 4 zusammengefaßt.

Eine gute Modellanpassung zeigt sich danach nur für den 2. Einzeltest. Daraus kann geschlossen werden, daß die Beantwortung der Fragebogenitems tatsächlich durch Antworttendenzen mitbestimmt wird. Ihre Wirksamkeit ist bei der Beantwortung der Fragen des 1. und 4. Tests am stärksten, bei der Beantwortung des 3. Tests noch deutlich ausgeprägt und beim 2. Test schwach bzw. vernachlässigbar. Eine Erklärung dafür könnte darin gesehen werden, daß die Art der Fragen, wie sie in den Fragebogen zu beantworten waren, bei der Beantwortung des 1. Einzeltests für die Schüler noch neu und ungewohnt war, während sie

beim 3. und 4. Test den Schülern bereits sehr geläufig war. Beide Bedingungen, sowohl die Ungewohntheit als auch die allzu große Geläufigkeit, machen möglicherweise die Fragen für eine "routinemäßige" Beantwortung anfällig.

Für die Messung von sachbezogener Motivation mit Hilfe von Fragebogen folgt daraus, daß die Antworten auf die positiv formulierten Fragen und die Antworten auf die negativ formulierten Fragen getrennt voneinander ausgewertet werden müssen. Die daraus geschätzten Meßwerte für die sachbezogene Motivation der Schüler sind durch die Antworttendenzen in entgegengesetzter Richtung verzerrt. Durch den Vergleich dieser verzerrten Meßwerte können dann sowohl die Ausprägung der Antworttendenzen als auch die Stärke der sachbezogenen Motiviertheit bestimmt werden.

#### 4. Diskussion

Mit der Konzeption der sachbezogenen Motivation im Schulunterricht als dem Ausmaß der Zuwendung des Schülers zu den im Unterricht vermittelten Lernstoffen, seiner Bereitschaft, sich mit diesen Sachverhalten freiwillig zu beschäftigen und weitere Informationen dazu aufzunehmen, sollte ein Motivationsbegriff präzisiert werden, der den Erfordernissen der motivationalen Evaluation von Schulunterricht gerecht wird. Zur Messung der sachbezogenen Motivation wurde ein Fragebogen konstruiert. Dabei wurde der Anspruch möglichst universeller Anwendbarkeit gestellt. Der Fragebogen sollte für verschiedene Schulstufen, Fächer, Inhalte, Methoden und Lehrer verwendbar sein und vergleichbare Werte liefern. Wie sich zeigte, kann dieser Anspruch jedoch nur bedingt aufrechterhalten werden. Das Konzept der sachbezogenen Motivation bedeutet eine gewisse Vereinfachung indem es nur einen Aspekt des als motiviert geltenden Verhaltens betrachtet und in den Vordergrund stellt. Die Verhaltensweisen, welche als Indiz für sachbezogene Motiviertheit gelten, können zugleich auch von anderen Bedingungen bestimmt werden. Die Zusammenfassung einer Klasse von Verhaltensweisen als Indiz für sachbezogene Motiviertheit und der Vergleich der Schüler aufgrund dieser Verhaltensweisen ist daher nur dann gerechtfertigt, wenn die Randbedingungen dieser Verhaltensweisen innerhalb der Gruppe der zu vergleichenden Schüler weitgehend dieselben bzw. in einem wohldefinierten Sinne äquivalent sind:

sie müssen das als Indiz für sachbezogene Motiviertheit herangezogene Verhalten (etwa) gleich stark begünstigen. Ist die Voraussetzung nicht erfüllt, so geht die Objektivität der Vergleich verloren.

Als eine Randbedingung, deren Konstanz nicht als selbstverständlich vorausgesetzt werden kann, erwiesen sich die "Lesegewohnheiten" der Schüler. In welcher Weise die Variabilität dieser Randbedingung zum Verlust der Objektivität führen kann, soll anhand eines Beispiels demonstriert werden: die Zustimmung zu der Feststellung, "Wenn mir jemand ein Buch über die zuletzt behandelten Themen schenken würde, würde ich mich darüber nicht besonders freuen", bedeutet nicht unbedingt ein mangelndes Interesse an diesen Themen, sondern kann auch dadurch bedingt sein, daß der Schüler generell wenig Freude an Büchern hat. Ist die Freude an Büchern bei den zu vergleichenden Schülern etwa gleich ausgeprägt, so wird sich daraus keine wesentliche Verzerrung der Aussagen über die sachbezogene Motiviertheit der Schüler ergeben. Findet sich dagegen unter den Schülern einer, der im Gegensatz zu den anderen sehr gerne liest, viel Freude an Büchern findet, so wird die sachbezogene Motiviertheit dieses Schülers im Vergleich zu den anderen überschätzt werden.

Wie sich gezeigt hat, kann aber auch der Unterricht selbst die Wirksamkeit von Randbedingungen mitbestimmen bzw. zu einer relativen Stabilisierung der Randbedingungen beitragen. Nach einem Gruppenunterricht, in dem die Schüler einen Großteil des Lernstoffs selbst aus der Literatur erarbeitet hatten, war der störende Einfluß der Lesegewohnheiten nicht mehr nachweisbar.

Die Evaluation von Unterricht nach dem Ausmaß der in den Schülern erweckten sachbezogenen Motiviertheit bleibt aber auch dann problematisch, wenn die Randbedingungen konstant sind. Sie geht nämlich von der impliziten Voraussetzung aus, daß der Unterricht zu all den Verhaltensweisen, die als Indiz für sachbezogene Motiviertheit herangezogen werden, gleichermaßen anregt, eben nur die Motiviertheit der Schüler verändert. Ist nach Inhalt und Methode sehr unterschiedlicher Unterricht zu evaluieren, so wird diese Voraussetzung aber nicht immer erfüllt sein. Es ist möglich, daß manche Unterrichtsformen bestimmte Verhaltensweisen stärker begünstigen als andere. Die "Schwierigkeiten" dieser Verhaltensweisen sind dann als Funktion des Unterrichts aufzufassen. In den oben diskutierten Itemanalysen wurde dies vor allem bei solchen Items deutlich, die die Zustimmung des Schülers zum Un-

terricht selbst erfragen und dabei eine ausgewogene Relation zwischen der Motiviertheit einerseits und dem im Unterricht realisierten Informationsangebot andererseits ansprechen. Allgemein gilt: Der Vergleich unterschiedlichen Unterrichts nach dem Ausmaß an sachbezogener Motivation, welches er in den Schülern wecken kann, ist nur dann gerechtfertigt, wenn die verschiedenen als motiviert geltenden Verhaltensweisen gleichermaßen angesprochen werden.

Evaluation von Unterricht im Hinblick auf sachbezogene Motiviertheit der Schüler muß daher dreierlei bedeuten:

- (1) Evaluation der Randbedingungen des motivierten Verhaltens: als Voraussetzung für die Objektivität.
- (2) Differentielle Evaluation der Motivierungsrichtung: wozu motiviert der Unterricht wie stark (Analyse der Schwierigkeiten der Verhaltensweisen als Funktion des Unterrichts).
- (3) Differentielle Evaluation der Motiviertheit: welche Schüler motiviert der Unterricht stärker als andere und warum?

Die Möglichkeit der Konstruktion universell anwendbarer Motivationstests scheidet damit aus. Stattdessen wird man anstreben, Itempools herzustellen, für die die Randbedingungen innerhalb definierter Schülerpopulationen und unter gegebenen Unterrichtsbedingungen hinreichend stabil sind und die unter den Unterrichtsbedingungen (etwa) gleichermaßen angesprochen werden. Für die Praxis bedeutet dies, daß der Geltungsbereich des der Messung der sachbezogenen Motiviertheit zugrundegelegten Meßmodells durch den zu evaluierenden Unterricht und die zu evaluierende Schülerpopulation immer wieder neu zu definieren ist. .

#### Literaturangaben

E. B. ANDERSEN: Das mehrkategoriale logistische Testmodell von RASCH. In: W. F. KEMPF (Ed.): Probabilistische Modelle in der Sozialpsychologie. Bern: Huber, 1974.

G. H. FISCHER: Neue Entwicklungen in der Psychologischen Testtheorie. In: G. H. FISCHER (Ed.): Psychologische Testtheorie. Bern: Huber, 1968.

- H. HECKHAUSEN: Leistungsmotivation. In: H. THOMAE (Ed.): Motivationslehre. Handbuch der Psychologie, Band 2. Göttingen: Hogrefe, 1965.
- H. HECKHAUSEN: Einleitung zu dem Kapitel "Motivation". In: W. EDELSTEIN und D. HOPF: Bedingungen des Bildungsprozesses. Stuttgart: Klett, 1973.
- W. F. KEMPF: Basisprobleme der Diagnostik der Aggressivität. In: L. ECKENSBERGER (Ed.): Bericht über den 28. Kongreß der Deutschen Gesellschaft für Psychologie. Göttingen: Hogrefe, 1974a.
- W. F. KEMPF: Dynamische Modelle zur Messung sozialer Verhaltensdispositionen. In: W. F. KEMPF (Ed.): Probabilistische Modelle in der Sozialpsychologie. Bern: Huber, 1974b.
- J. V. MITCHELL: An Analysis of the Factorial Dimensions of the Achievement Motivation Construct. In: Journal of Educational Psychology 52 (1961), S. 179-187.
- P. H. PEARSON: Relationships between global and specified measures of novelty seeking. In: Journal Cons. Clin. Psychology 34 (1970), S. 199-204.
- G. RASCH: Probabilistic models for some intelligence and attainment tests. Copenhagen: Danmarks Paedagogiske Institut, 1960.
- J. VEROFF: Wie allgemein ist das Leistungsmotiv? In: W. EDELSTEIN und D. HOPF: Bedingungen des Bildungsprozesses. Stuttgart: Klett, 1973.