

Bellack, Arno A.

Methoden zur Beobachtung des Unterrichtsverhaltens von Lehrern und Schülern

Wulf, Christoph [Hrsg.]: *Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. München : R. Piper & Co. Verlag 1972, S. 211-238. - (Erziehung in Wissenschaft und Praxis; 18)



Quellenangabe/ Reference:

Bellack, Arno A.: Methoden zur Beobachtung des Unterrichtsverhaltens von Lehrern und Schülern - In: Wulf, Christoph [Hrsg.]: *Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen*. München : R. Piper & Co. Verlag 1972, S. 211-238 - URN: urn:nbn:de:0111-opus-14281 - DOI: 10.25656/01:1428

<https://nbn-resolving.org/urn:nbn:de:0111-opus-14281>

<https://doi.org/10.25656/01:1428>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, veröffentlichen oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Digitalisiert

Mitglied der


Leibniz-Gemeinschaft

Evaluation

Beschreibung und Bewertung von Unterricht,
Curricula und Schulversuchen

Texte

herausgegeben von Christoph Wulf



R. Piper & Co. Verlag
München

ISBN 3-492-01985-4
© R. Piper & Co. Verlag, München 1972
Gesamtherstellung Clausen & Bosse, Leck/Schleswig
Umschlagentwurf Gerhard M. Hotop
Printed in Germany

ARNO A. BELLACK

*Methoden zur Beobachtung des Unterrichtsverhaltens
von Lehrern und Schülern*

Während der letzten zehn Jahre wurde durch das aufkommende Interesse an der Erforschung des Unterrichtsverhaltens eine bedeutsame Entwicklung in der pädagogischen Forschung in Gang gesetzt. Es handelt sich hierbei um das Wiederaufgreifen eines seit langem existierenden Anliegens der Erziehungswissenschaftler. Bei den gegenwärtigen Unterrichtsforschungen, die sich von den Arbeiten vorangegangener Perioden unterscheiden, wird auf systematische Verhaltensbeobachtung von Schülern und Lehrern im Unterricht großer Wert gelegt.

Dieser Beitrag stellt eine Auseinandersetzung mit der neueren Entwicklung von Methoden der Verhaltensbeobachtung im Unterricht dar. Er gliedert sich in drei Abschnitte:

- I. Entwicklung von Systemen zur Verhaltensbeobachtung im Unterricht
- II. Beispiele für Beobachtungssysteme
- III. Probleme und Konsequenzen in der Entwicklung und im Gebrauch von Systemen zur Analyse des Lehrer- und Schülerverhaltens.

*I. Entwicklung von Systemen zur Verhaltensbeobachtung
im Unterricht*

Bei der Konstruktion eines Systems zur Beobachtung des Unterrichtsverhaltens muß der Forscher zwei fundamentale Fragen beantworten: (1) Welche Dimensionen des Unterrichtsverhaltens sollen beobachtet werden? (2) Auf welche Art und Weise sollen die Beobachtungen durchgeführt werden? Diese beiden Fragen sind natürlich eng miteinander verknüpft, denn Beobachtungsgegenstand und Beobachtungsmethode sind voneinander abhängig. Der Prozeß der Unterrichtsbeobachtung wird von den Konzeptionen des Forschers über das Verhalten, das er untersucht, und von den Verfahrensweisen bestimmt, mit denen er die ihm zur Verfügung stehenden Daten sammelt.

Welches Unterrichtsverhalten soll beobachtet werden?

Da der Forscher keinesfalls alles in der Klasse beobachten kann, muß er eine Entscheidung darüber fällen, welche Aspekte des Lehrer- und Schülerverhaltens er zum Gegenstand seiner Untersuchungen nimmt. »Jede Beobachtung wird gemacht«, schreibt Kaplan, »sie ist das Produkt einer aktiven Wahl, nicht eines passiven Geschehens. Keine einzige Interpretation ist notwendigerweise die Folge von dem, was beobachtet wurde; es gibt immer viele Möglichkeiten, Verhalten zu klassifizieren« (Kaplan 1964, 133).

Wie der Forscher bei der Klassifikation des Unterrichtsverhaltens verfährt, wird vom Ziel seiner Untersuchung bestimmt, ebenso vom konzeptuellen Rahmen, der den Beobachtungsgegenstand, die Beobachtungsmethode und den Umfang der Beobachtung angibt. Flanders (1965) z. B. entwickelte auf Grund seines theoretischen Interesses am Klassenklima eine Reihe von Kategorien, um die verschiedenartigen Einflüsse des verbalen Lehrerverhaltens zu beschreiben. Im Gegensatz dazu konstruierten B. O. Smith und Meux (1962) auf Grund ihres Interesses an den kognitiven Dimensionen des Unterrichtsgeschehens einen Katalog von Beobachtungskriterien, um logische Operationen im Unterricht zu kategorisieren. Kounin (1967), dessen Interesse sich auf Klassenführung und Disziplin konzentrierte, entwickelte ein Instrument zur Klassifikation von Dimensionen des Lehrerstils, wie er sich bei der Kontrolle des Schülerverhaltens zeigt.

Zur systematischen Beobachtung des Unterrichtsverhaltens, das für die Zielsetzung und den theoretischen Anspruch relevant ist, muß der Wissenschaftler ein zuverlässiges und gültiges Instrument in Form von Schätzskaleten (rating scales) bzw. ein Beobachtungssystem entwickeln, das in operationalisierter Form die zu kategorisierenden oder zu messenden Verhaltensdimensionen spezifiziert.

Schätzskaleten: Beim Gebrauch von Schätzskaleten besteht die Aufgabe des Beobachters darin, das beobachtete Verhalten an einem Ort auf einem Kontinuum oder in einem in einer bestimmten Reihenfolge vorgegebenen Kategoriensystem zu lokalisieren (vgl. Remmers 1963). Ein Beispiel für eine experimentell entwickelte Schätzskaleten ist die von Ryan (1960) für seine Erforschung von Lehrercharakteristika konstruierte Skala. In dieser Untersuchung bewertete der Beobachter, nachdem er einer Unterrichtsstunde beigewohnt hatte, jede der sechsundzwanzig Dimensionen des während des Unterrichts beobachteten Verhaltens und übertrug anschließend seine Bewertung in Form von Schätzungen auf 7-Punkte-Skaleten. Die 22 Dimensionen des Lehrerverhaltens bezogen sich u. a. auf folgende Kategorien: parteiisch – fair; barsch – freundlich; zurück-

haltend – temperamentvoll; autokratisch – demokratisch und unsicher – sicher. Zu den Dimensionen des Schülerverhaltens gehörten: teilnahmslos – lebhaft; sich auf andere verlassend – die Initiative selbst ergreifend; Widerstand leistend – sich verantwortlich fühlend. Jede Dimension war in einem speziellen Wörterverzeichnis definiert, wobei die relevanten Verhaltensaspekte, auf denen die Schätzung basierte, gesondert angegeben wurden.

Die Verfahren, in denen die Schätzung *nach* der Beobachtung erfolgte – wie bei den von Ryan (1960) entwickelten – unterliegen allerdings bedeutsamen Einschränkungen, wenn mit ihrer Hilfe die Aktivitäten im Unterricht erfaßt werden sollen. Die Einschätzungen durch die Beurteiler liefern lediglich allgemeine Eindrücke und Erinnerungen an das tatsächliche Unterrichtsgeschehen, nicht jedoch exakte Aufzeichnungen des Lehrer- und Schülerverhaltens. Deshalb und wegen gewisser Unzulänglichkeiten der Schätzskalen als Meßinstrumente werden sie nur in wenigen gegenwärtigen Untersuchungen verwendet, die die direkte Beobachtung des Unterrichtsgeschehens zum Gegenstand haben, obwohl sie in früheren Untersuchungen häufig Anwendung fanden.

Beobachtungssysteme: Die Systeme zur Beobachtung des Unterrichtsverhaltens liefern dem Beobachter eine Reihe von Kategorien, denen das jeweilige Verhalten zugeordnet wird. Meistens sind Verhaltensstimuli oder operationale Definitionen der Kategorien und Kodierungsanweisungen vorhanden, um dem Beobachter die Entscheidung zu erleichtern, welcher der Kategorien das beobachtete Verhalten zuzuordnen ist. Während der letzten Jahre wurde eine große Anzahl von Beobachtungssystemen entwickelt. Simon und Boyer (1967) katalogisierten in ihrer Anthologie von Beobachtungsinstrumenten, »Mirrors for Behavior«, 26 Verfahren; doch enthält sie lediglich die Hälfte aller in der letzten Zeit entwickelten Systeme¹.

Zwischen den vorhandenen Beobachtungssystemen zeigen sich große Unterschiede. Es bestehen Differenzen

1. in den Dimensionen des zu klassifizierenden Unterrichtsverhaltens,
2. in der Art des zur Beobachtung entwickelten Kriterienkatalogs,
3. in dem Bezugsrahmen des Beobachters für das Kodieren,
4. in der beim Kodieren zu benutzenden Verhaltenseinheit,
5. im Anwendungsbereich.

Diese Unterschiede verdeutlichen die Grundproblematik bei der Konstruktion von Beobachtungsinstrumenten.

1. Dimensionen des Unterrichtsverhaltens

Wie bereits dargelegt, werden bei den vorliegenden Systemen zur Verhaltensanalyse bestimmte Aspekte besonders hervorgehoben, andere vernachlässigt. Die vorhandenen Beobachtungskataloge enthalten ein weites Spektrum von Verhaltensweisen. Biddle (1967) fand bei einer Aufarbeitung der Forschungen über Unterrichtsverhalten, daß kürzlich entwickelte Instrumente sich befassen

- a) mit dem Lehrerverhalten im Sinne von Handlungsweisen, Methoden und charakteristischen Rollen,
- b) mit dem Zuhörer- und Zielverhalten von Lehrern und Schülern,
- c) mit der Lehrer-Schüler-Interaktion,
- d) mit äußeren Strukturen – wie Lehrinhalten und administrativen Maßnahmen,
- e) mit inneren Strukturen – wie Kommunikationsstruktur, Aktivitätsstruktur, charakteristischen Rollen, sozialen Funktionen.

Im Gegensatz dazu haben Simon und Boyer (1967) verschiedene Beobachtungsverfahren in drei Gruppen nach folgender Systematik kategorisiert:

- a) affektive Systeme, mit deren Hilfe man das emotionale Klassenklima erfassen kann und die Art und Weise, wie es durch die Reaktion der Lehrer auf die Emotionen, Ideen oder Handlungsweisen der Schüler bedingt wird,
- b) kognitive Systeme, die sich mit Denkprozessen befassen und den diese zum Ausdruck bringenden verbalen Verhaltensmustern,
- c) multidimensionale Systeme, mit denen man sowohl die kognitiven als auch die affektiven Dimensionen des Verhaltens erfassen kann.

Das Flanderssche System der Interaktionsanalyse (1965) und das Hughessche System zur Klassifizierung der Funktionen des Lehrerverhaltens (1959) sind bedeutsame Beispiele für das erste System. Das System von Bellack und anderen zur Beschreibung des Sprachverhaltens im Unterricht (1966) und B. O. Smiths' und Meux' Verfahren zur Analyse logischer Operationen des Unterrichtens (1962) sind typisch für die zweite Gruppe. Olivers und Shavers Beobachtungssystem zur Beschreibung kontrastierender Unterrichtsstile in dem Fach Sozialkunde (social studies) (1966) sowie das System von Joyce zur Erfassung affektiver und kognitiver Aspekte der verbalen Lehrerkommunikation (vgl. Joyce/Harootunian 1967) sind Repräsentanten der dritten Kategorie.

Die vorhandenen Systeme zur Beobachtung des Unterrichtsverhaltens können ebenso nach der Art der Kommunikation klassifiziert werden, d. h. danach, ob sie die Aufmerksamkeit auf (a) verbale Aspekte, (b) nicht-ver-

bale Aspekte oder (c) verbale und nicht-verbale Aspekte der Kommunikation konzentrieren. Die meisten Systeme richten sich ausschließlich auf verbales Verhalten, während einige Systeme verbale und nicht-verbale Dimensionen beinhalten. Zwanzig der Systeme, die in »Mirrors for Behavior« (Simon und Boyer 1967) beschrieben sind, befassen sich mit verbaler Kommunikation, nur sechs beziehen sich sowohl auf verbales als auch auf nicht-verbales Verhalten. Soweit der Autor informiert ist, ist Galloway (1962) der einzige Forscher, der für seine Untersuchungen über nicht-verbale Kommunikation, wie die Gestik des Lehrers, den Ausdruck der Stimme und die Mimik, ein Instrument nur zur Beschreibung des nicht-verbalen Verhaltens entwickelt hat.

Die Entscheidung, welche Aspekte des Lehrer-Schüler-Verhaltens zu untersuchen sind, ist prinzipiell ein theoretisches Problem. Es sind jedoch auch wichtige methodische Erwägungen miteinbezogen, z. B. ist dafür zu sorgen, daß die Kategorien in einem vorgegebenen System sich gegenseitig ausschließen und einen definierten Bereich erschöpfend repräsentieren. Das Kriterium der Vollständigkeit bedeutet natürlich nicht, daß jedes Verhalten während einer bestimmten Beobachtungsperiode klassifiziert wird. Vielmehr ist es erforderlich, daß das zu beobachtende Verhaltensuniversum (z. B. soziales Klima, logische Operationen, Kommunikationsprozesse) klar definiert und eine repräsentative Stichprobe von Kategorien aus diesem Verhaltensuniversum entnommen wird. Zum Beispiel schließen sich die folgenden fünf von Aschner und Gallagher (1965) entwickelten allgemeinen Kategorien zur Klassifikation von Gedankenprozessen, die sich im verbalen Unterrichtsverhalten widerspiegeln, gegenseitig aus und sind auch repräsentativ für ein Universum kognitiver Funktionen: mechanisches Denken, Gedächtnistätigkeit, konvergentes Denken, evaluatives Denken, divergentes Denken.

In dem Stadium der Forschungsarbeit, in dem die relevanten theoretischen Konzepte noch nicht vollständig entwickelt sind, ist es häufig schwer, Kategorien aufzustellen, die sich gegenseitig ausschließen und die repräsentativ sind. Jedoch muß der Wissenschaftler im Verlauf seiner Forschungsarbeit diese Kriterien ständig beachten und sich nicht nur auf die Analyse der empirischen Beobachtungsdaten, sondern ebenso auf die Verbesserung der zur Analyse der Daten verwendeten Begriffe konzentrieren.

2. Verschiedene Arten von Beobachtungssystemen

In der Erforschung des Unterrichtsverhaltens werden heute weitgehend zwei Arten von Beobachtungssystemen angewandt; Medley und Mitzel (1963) bezeichnen diese als Kategoriensysteme und Zeichensysteme.

Ein Kategoriensystem beschränkt die Beobachtung auf spezifische Verhaltensdimensionen im Unterricht, wobei es eine Reihe von Kategorien liefert, in die jede beobachtete Verhaltenseinheit eingestuft wird. Das daraus resultierende Ergebnis zeigt für jeden Beobachtungszeitraum die vollständige Anzahl der aufgetretenen Verhaltenseinheiten an sowie die Anzahl innerhalb jeder einzelnen Kategorie. Die von Hughes (1959) in ihrer Untersuchung über die Interaktion in der Primarstufe entwickelte Reihe von Kategorien ist ein typisches Beispiel dafür. Der dieses System benutzende Beobachter klassifiziert das verbale und nicht-verbale Lehrerverhalten nach sechs Hauptfunktionen: Kontrollfunktionen, unterstützende Funktionen, Funktionen, die die Vermittlung von Unterrichtsinhalten zum Gegenstand haben, Funktionen, die als persönliche Antwort dienen, Funktionen von positivem und von negativem Gefühlswert. Im Gegensatz dazu liefert ein Zeichensystem dem Beobachter eine Reihe spezifischer Verhaltensweisen, die während einer Beobachtungsperiode auftreten oder nicht auftreten können. Die Beobachter werden angewiesen, auf diese spezifischen Verhaltensweisen zu achten, und das Ergebnis zeigt, welche dieser Verhaltensweisen während des Beobachtungszeitraums aufgetreten sind. Ein Beispiel für ein solches Beobachtungssystem ist die von Medley und Mitzel (1958) entwickelte Beobachtungsskala (OSCAR). Ein Beobachter, der dieses Instrument verwendet, registriert lediglich die Verhaltensweisen, die von den 71 Items erfaßt werden. Zu diesen in mehrere Abschnitte gruppierten Items gehören etwa solche: »der Lehrer trägt vor«, »der Lehrer beantwortet die Fragen der Schüler«, »der Lehrer veranschaulicht etwas an der Tafel«, »der Lehrer reagiert sarkastisch«, »der Schüler spricht zur Gruppe«, »der Schüler flüstert«, »der Schüler liest oder arbeitet an seinem Platz«.

Während Zeichensysteme gewöhnlich aus einer großen Itemanzahl bestehen, die sich auf konkrete, spezifische Verhaltensweisen beziehen und daher ein geringes Maß an Schlußfolgerungen seitens des Beobachters erfordern, sind Kategoriensysteme gewöhnlich aus einer geringeren Anzahl von Items zusammengestellt und besitzen ein höheres Abstraktionsniveau, das einen höheren Grad an schlußfolgerndem Denken beim Beobachter voraussetzt. Nicht alle Kategoriensysteme sind auf dem gleichen Abstraktionsniveau, aber in der Regel auf einem höheren als die Zeichensysteme konstruiert. Die in den beiden vorhergehenden Abschnitten erwähnten Beispiele veranschaulichen den Unterschied zwischen mehr oder weniger schlußfolgerndes Denken erfordernden Kategorien. Der Beobachter ist weniger auf Schlußfolgerungen angewiesen, wenn er die Handlung eines Lehrers zu interpretieren hat, wie z. B. »Beantwortung der Schülerfragen« oder »Veranschaulichung an der Tafel«. Im Gegensatz dazu er-

fordert es einen höheren Grad an schlußfolgerndem Denken, eine Aussage des Lehrers als Kontroll- oder Unterstützungsfunktion zu interpretieren. Es sollte ebenso festgehalten werden, daß Kategorien, die in einem hohen Maße schlußfolgerndes Denken erfordern und die die Verhaltensbeobachtung als Basis für Schlußfolgerungen über Motive oder Auswirkungen des Verhaltens beinhalten, den Forscher bei der Validitätsbestimmung mit schwierigeren Problemen konfrontieren, als es bei Kategorien der Fall ist, die sich mit deskriptiven Verhaltensbeobachtungen befassen und oft als solche bezeichnet werden, die Augenscheinvalidität besitzen.

Medley und Mitzel (1963) stellen fest, daß Kategoriensysteme häufiger in empirischen Untersuchungen angewandt werden, die auf hochdifferenzierten Theorien beruhen, während Zeichensysteme dann verwendet werden, wenn die Theorie nicht genügend aussagekräftig ist. Es ist selbstverständlich, daß die theoretische Orientierung des Forschers und der Grad der Entwicklung seiner Theorie das Niveau der Konzeptualisierung und folglich auch den Grad des erforderlichen Maßes an schlußfolgerndem Denken bestimmen.

3. Bezugsrahmen des Beobachters

Die Wissenschaftler versuchen, die Vorgänge im Schulunterricht wenigstens unter drei Aspekten zu kategorisieren:

- a) die Absicht oder das Motiv des Handelnden
- b) die Auswirkungen des Verhaltens auf den Adressaten
- c) die objektiven Verhaltenscharakteristika.

Beispiele dieser drei Aspekte finden sich in neueren Untersuchungen über Unterrichtsprozesse.

Das analytische System von Withall (1949), das er in Verbindung mit seinen Forschungen über sozial-emotionales Klassenklima entwickelt hat, erfordert vom Beobachter eine Interpretation der Aussagen des Lehrers dahingehend, ob ihre Intention eine schülerunterstützende, problemstrukturierende oder direktive ist. Hughes (1959) andererseits klassifizierte in ihrer Untersuchung über das Unterrichtsverhalten von Primarschullehrern die Funktionen des verbalen und nicht-verbalen Lehrerverhaltens nach ihren erwarteten Wirkungen und deren Bedeutung für die Schüler. Im Gegensatz dazu beschreiben Smith und Meux (1962) in ihrer Untersuchung über die logischen Aspekte des Unterrichtsgeschehens die logischen Bestandteile der Lehrer- und Schüleraussagen, wobei sie über Motive des Sprechers oder seine Wirkungen auf die Zuhörer keine Aussagen machen.

Die Erfahrungen dieser und anderer Forscher verdeutlichen, daß das Unterrichtsverhalten auf der Basis aller drei Aspekte zuverlässig kodiert werden kann. Welcher der drei Aspekte für eine Untersuchung angemessen

sen ist, kann nur entschieden werden, wenn man das Ziel der Untersuchung berücksichtigt. Biddle (1967) hat mehrere mögliche Zielsetzungen genannt, die für die Unterrichtsforschung in Frage kommen, sowie Konzeptionen entwickelt, die diesen Zielen angemessen sind. Zum Beispiel ist er der Auffassung, daß Urteile über die Intention des Lehrers angemessen sind, wenn man primär an den Determinanten des Lehrerverhaltens interessiert ist. Wenn man sich allerdings mit der Qualifikation des Lehrers befaßt, sind Urteile über die Wirkung des Lehrerverhaltens auf die Lernprozesse der Schüler angebrachter. Wenn andererseits die individuellen und sozialen Determinanten des Verhaltens erforscht werden sollen oder kontrastierende Modelle der Interaktion im Klassenraum zu testen sind, wäre es sinnvoll, objektive Verhaltenscharakteristika als Grundlage zu nehmen.

4. Verhaltenseinheiten

Eine entscheidende Aufgabe bei der Erstellung eines Beobachtungssystems besteht darin, die Verhaltenseinheit zu spezifizieren, die als Basis für das Kodieren benutzt wird. Für die vielen Möglichkeiten, die Verhaltenseinheit zu definieren, gibt es zwei grundsätzlich verschiedene Methoden: (a) Bestimmung einer willkürlichen Zeiteinheit, (b) genaue Beschreibung einer analytischen Einheit, die häufig durch das Kategoriensystem selbst vorgegeben wird.

Bei den Systemen, bei denen eine willkürliche Zeiteinheit benutzt wird, muß der Beobachter ein Protokoll über die in dieser festgesetzten Zeit auftretenden Verhaltensweisen erstellen. Flanders (1965) z. B. verlangt vom Beobachter alle drei Sekunden ein Urteil darüber, ob der Lehrer innerhalb der kurzen Periode direkten oder indirekten Einfluß auf die Schüler ausübt. In gleicher Weise schreibt Spaulding (vgl. Simon/Boyer 1967) Zeitperioden von 3–10 Sekunden vor, für die der Beobachter ein Urteil über das Verhalten des Lehrers hinsichtlich seiner Kontrollfunktion in der Klasse abgibt, wobei er Kategorien wie »Setzen von Verhaltenszielen« und »Anordnen bestimmter Aktivitäten« verwendet. Medley und Mitzel (1958) andererseits erwarten, daß der Beobachter Verhaltensweisen in 5-Minuten-Perioden protokolliert.

Der grundlegende Vorteil der Setzung beliebiger Zeiteinheiten liegt in ihrem mechanischen Charakter, der eine Hilfe für die Regulierung des Beobachtungsprozesses darstellt. Weiterhin kann diese Methode bei der Analyse des verbalen und des nicht-verbalen Verhaltens angewandt werden. Jedoch liegt die Problematik der Zerlegung des Unterrichtsgeschehens in willkürliche Zeiteinheiten darin, daß die Ergebnisse nicht die na-

türlich eintretenden Verhaltensmuster oder den Ablauf des Verhaltens widerspiegeln, wie es sich allmählich entwickelt.

Anstatt eine beliebige Zeiteinheit anzugeben, entscheiden sich viele Forscher für Analyse-Einheiten verschiedener Art. Diese Einheiten repräsentieren einzelne Elemente des verbalen und/oder nicht-verbalen Verhaltens, die beim Kodieren verwendet werden, um Unterrichtsgeschehnisse in Einzelkomponenten zu zerlegen. Analyse-Einheiten werden von Wissenschaftlern unterschiedlich definiert, und zwar im Sinne von:

- a) Aktivität im Unterricht,
- b) verbaler Kommunikation zwischen zwei oder mehreren Sprechern,
- c) Kommunikation oder Mitteilung eines einzelnen Sprechers,
- d) Ausdruck einer Gedankeneinheit durch einen Sprecher.

Diese Analyse-Einheiten können durch folgende Beispiele veranschaulicht werden:

a) In seiner Untersuchung über Unterrichtsverhalten und mangelndes Leistungsvermögen der Schüler führte Perkins (1964, 1965) als grundlegende Analyse-Einheiten sechs Arten von Klassenaktivitäten auf: Großgruppendifkussion, Klassenvortrag, Einzel- oder Projektarbeit, die nicht allen Schülern zugewiesen ist, Bearbeitung einer allen zugewiesenen Aufgabe, Kleingruppenarbeit, mündliche Berichterstattung. Aspekte des Schüler- und Lehrerverhaltens wurden im Kontext mit diesen Aktivitätsarten kodiert.

b) In ihrer Untersuchung der logischen Dimensionen des Unterrichtsverlaufs bezeichneten B. O. Smith und Meux (1962) als fundamentale Verhaltenseinheit den verbalen Austausch zwischen zwei oder mehreren Sprechern, wobei drei Phasen zu unterscheiden sind: Anfangs- oder Eröffnungsphase, Durchführungsphase und Endphase.

c) Jackson (1965) definierte drei Arten von verbalen Äußerungen als fundamentale Einheiten zur Beschreibung der Kommunikation in der Primarstufe: Äußerungen, die sich auf (a) inhaltliche Ziele, (b) Verfahrensweisen und Regeln bei der Gruppenarbeit, (c) Mitteilungen zur Aufrechterhaltung der Disziplin und Ordnung beziehen.

d) In ihrer Untersuchung über die Lehrstrategien zur Schulung kognitiver Fähigkeiten bei Primarschülern bestimmten Taba u. a. (1964, 1966) die fundamentale Analyse-Einheit als »Gedanken-Einheit«, definiert als »Äußerung oder Serie von Äußerungen, die einen mehr oder weniger vollständigen Gedanken zum Ausdruck bringen, eine spezifische Funktion besitzen und entsprechend dem jeweiligen Niveau des Gedankens klassifiziert werden können« (1966, 134).

5. Anwendungsbereich

Die Beobachtungssysteme unterscheiden sich auch dadurch voneinander, daß sie in verschiedenem Ausmaß auf andere Forschungsvorhaben und Populationen anwendbar sind.

Einige Systeme wurden in der Absicht entwickelt, sie für verschiedene Klassen, Altersstufen, Fächer und Schülergruppen verwenden zu können. Zum Beispiel ist das System der Interaktionsanalyse von Flanders in Primar- und Sekundarstufen anwendbar, in denen Lehrer mit Schülern verbal in Interaktion treten. Gleichermaßen wurde das Verfahren zur Analyse der strukturellen und funktionellen Aspekte des Kommunikationssystems im Unterricht von Biddle und Adam bewußt für die Anwendung in Primar- und Sekundarstufen entwickelt, in denen zahlreiche Fächer unterrichtet werden.

Im Gegensatz dazu sind einige Beobachtungssysteme zur Analyse von Unterrichtsprozessen nur für bestimmte Altersstufen geeignet. Das System von Hughes (1959) zur Beschreibung der Lehrerfunktionen ist in Primarstufen anwendbar, während das System zur Analyse logischer Operationen des Unterrichtens von B. O. Smith und Meux (1962) lediglich für Klassen der Sekundarstufe in Frage kommt, in denen Fächer mit stärkerem Wissenschaftscharakter unterrichtet werden.

Andere Systeme sind auf bestimmte Fächer oder Schülergruppen beschränkt. Das System von Wright (1959) wurde zur Untersuchung des Mathematikunterrichts entwickelt, während das Verfahren von Oliver und Shaver (1966) vornehmlich zur Erforschung des Sozialkundeunterrichts (social studies) eingesetzt werden kann.

Einige Forscher befassen sich nur mit bestimmten Schülergruppen: Kounin (1967) mit emotional gestörten Schülern, Perkins (1964, 1965) mit leistungsschwachen Schülern und Smith und Geoffrey (1968) mit Schülern von sozioökonomisch niedrigerem Status.

Ob die Forscher versuchen sollten, umfassende multidimensionale Systeme zu konstruieren, die auf einen großen Bereich von Unterrichtssituationen anwendbar sind, oder sich darauf beschränken sollten, Instrumente für eine begrenzte Anzahl von Unterrichtssituationen zu entwickeln, ist eine umstrittene Frage, auf die später näher eingegangen wird.

Wie werden Unterrichtsbeobachtungen durchgeführt?

Registrieren und Kodieren von Verhaltensweisen

Eine Vielzahl verschiedener Registrier- und Kodieretechniken zur Fixierung von Verhaltensweisen wurde im Zusammenhang mit den oben erörterten

Beobachtungssystemen angewandt. Einige Forscher (z. B. Flanders 1965; Jackson 1965; Medley/Mitzel 1958) fordern vom Beobachter, das Verhalten so zu kodieren, wie es sich tatsächlich ereignet hat; d. h. der Beobachter, dem die Aufgabe des Kodierens zukommt, transformiert die beobachteten Ereignisse in Symbole, die gezählt und tabuliert werden können. Der Hauptvorteil beim unmittelbaren Kodieren des Verhaltens liegt darin, daß man direkten Zugang hat zu visuellen Reizen wie Mimik und Gestik der Lehrer und Schüler und zu situationsgebundenen Faktoren während des Unterrichts, die für die genaue Interpretation des beobachteten Verhaltens relevant sein können. Ein schwerwiegender Nachteil liegt darin, daß es äußerst schwierig ist, die komplexen Verhaltensweisen unmittelbar in dem Augenblick zuverlässig zu kodieren, in dem sie während des Unterrichts auftreten. Daher protokollieren viele Forscher zunächst das Verhalten während des Unterrichts und kodieren es anschließend mit Hilfe des Protokolls.

Bei der Erforschung des Unterrichtsverhaltens wurden verschiedene Arten der Berichterstattung verwendet, u. a. schriftliche Protokolle, Tonbandaufnahmen und audiovisuelle Aufzeichnungen. Hughes (1959) sammelte in ihrer Untersuchung über den Unterricht in der Primarstufe Daten in Form von schriftlichen Protokollen, die in der Hauptsache aus einem chronologischen Bericht über das verbale und nicht-verbale Lehrerverhalten im Unterricht bestanden. Zwei geübte Beobachter protokollierten gleichzeitig das Unterrichtsgeschehen; das endgültige Protokoll über eine Beobachtungsperiode bestand lediglich aus Beschreibungen, denen beide Beobachter zustimmten. Es wurde kein Versuch unternommen, das Verhalten zu kategorisieren; die Beobachter berichteten lediglich in Kurzfassung über die Aussagen und Handlungen der Lehrer. Beim Kodieren klassifizierte man später die protokollierten Daten in sieben Hauptkategorien hinsichtlich der Lehrerfunktionen.

Hughes vertritt die Auffassung, daß das Sammeln von Unterrichtsdaten in Form von schriftlichen Protokollen verschiedene Vorteile hat, z. B. ermöglicht dieses Vorgehen den Forschern, den Unterrichtsverlauf so zu fixieren, daß er hinsichtlich seiner spezifischen Qualität sowie seiner charakteristischen Merkmale untersucht werden kann; die Kontinuität des Lehrerverhaltens wird festgehalten; die Protokolle sind insofern neutral, als während der Beobachtung des Unterrichtsgeschehens keine Bewertung abgegeben wird. Diese Neutralität der schriftlichen Berichte muß jedoch ernsthaft in Frage gestellt werden, da die Protokolle über das Unterrichtsverhalten aus zweiter Hand durch den Beobachter abgefaßt werden, dessen Subjektivität der Wahrnehmung sich unvermeidbar in seinem Bericht über das Geschehen widerspiegelt.

Diese Unzulänglichkeit von schriftlichen Berichten kann weitgehend überwunden werden durch den Gebrauch von Tonbändern oder audiovisuellen Aufzeichnungen, die fortlaufende, objektive Verhaltensprotokolle liefern. In mehreren Untersuchungen zur Erforschung des verbalen Unterrichtsverhaltens (z. B. Bellack u. a. 1966; Taba u. a. 1964; Taba 1966; und B. O. Smith/Meux 1962) sind die Ereignisse mit Hilfe von Fernsehkameras aufgezeichnet worden. Häufig werden von Tonbandaufzeichnungen maschinenschriftliche Protokolle angefertigt; derjenige, dem die Aufgabe des Kodierens zufällt, hat bei der Kategorisierung des Verhaltens Zugang zu den Tonbandaufzeichnungen und den maschinenschriftlichen Protokollen. In Anbetracht der Schwierigkeit, das komplexe verbale Verhalten im Unterricht der Realität entsprechend zu beobachten und zu analysieren, haben diese Aufzeichnungen über den Unterrichtsverlauf offensichtlich Vorteile, da sie wiederholt von verschiedenen Perspektiven her analysiert werden können. Die prinzipielle Unzulänglichkeit von Tonbandaufnahmen liegt darin, daß sie keine Informationen über nicht-verbales Verhalten liefern, das als solches sehr bedeutsam sein oder als Hinweis auf die adäquate Interpretation des verbalen Verhaltens dienen kann.

Dieser Schwierigkeit kann durch die Verwendung von audiovisuellen Aufzeichnungen begegnet werden, die ein umfassendes Protokoll der Unterrichtsabläufe liefern, da sie verbales und nicht-verbales Verhalten erfassen. Kounin (1967) und Biddle und Adams (1967) haben Untersuchungen abgeschlossen, in denen Beobachtungsdaten während des regulären Unterrichts mit transportablen Videorekordern aufgezeichnet wurden. Die Hauptprobleme liegen in den damit verbundenen Kosten und den technischen Schwierigkeiten beim Gebrauch der komplizierten Apparatur. Der Leser, der am Einsatz von Videorekordern in der Unterrichtsforschung interessiert ist, wird auf den Bericht von Biddle und Adams verwiesen (1967).

Beziehungen zwischen Beobachter und Beobachteten

Ein häufig erhobener Einwand gegen Beobachtungsverfahren liegt darin, daß die Gegenwart eines Beobachters oder eines Aufnahmeinstruments derart ablenkend ist, daß das beobachtete Verhalten nicht als typisches Verhalten angesehen werden kann. Dies scheint jedoch nach den Urteilen erfahrener Forscher keine stichhaltige Kritik zu sein. Heyns und Lippitt stellen in ihrem umfangreichen Überblick über systematische Beobachtungstechniken in der sozialpsychologischen Forschung fest, daß die Wissenschaftler, die Beobachter zur Erforschung des Unterrichtsgeschehens einsetzen, die allgemeine Auffassung teilen, daß »die Beobach-

ter, wenn überhaupt, nur sehr geringe Wirkung ausüben. Dieser Überzeugung sind auch experimentell arbeitende Forscher, die in einer Vielfalt von Situationen und auf verschiedenartigen Fachgebieten arbeiten« (Heyns/Lippitt 1954, 399).

Die meisten Forscher im Bereich der Erziehungswissenschaft scheinen diese Auffassung zu teilen. Biddle und Adams (1967), deren komplizierte audiovisuelle Apparatur aus zwei Kameras und zwei Mikrofonen bestand, die in den Klassenräumen aufgestellt wurden, in denen sie ihre Beobachtungen vornahmen, kommentieren die Auswirkungen dieser Apparatur auf Schüler und Lehrer folgendermaßen:

»Alle beteiligten Lehrer wurden informell über den Effekt der Aufnahmeapparatur interviewt. Einige berichteten über gewisse Spannungsgefühle zu Beginn der ersten protokollierten Unterrichtseinheit, stellten jedoch fest, daß diese Spannung dann aufhörte, wenn sie sich im Unterricht engagierten. Die Forscher bemerkten, daß einige Lehrer sich an den Tagen besonders sorgfältig zu kleiden schienen, an denen die Aufnahmen stattfanden, obwohl ein derartiger Effekt bei den Schülern nicht festgestellt wurde, die – es soll noch einmal daran erinnert werden – nicht wußten, an welchen Tagen die Aufnahmen gemacht wurden. Gelegentlich zeigten sich deutliche Anzeichen dafür, daß die Schüler sich der Aufnahmekameras bewußt waren, da sie auf die vermuteten Kameras starrten oder in der Pause vor ihnen Theater spielten. Jedoch wurde überall bestätigt, daß die Kameras die Schüler offensichtlich kaum ablenkten« (Biddle/Adams 1967, 217).

Ähnliche Auffassungen, daß die von Beobachter und Aufnahmegerät ausgehenden Wirkungen auf das Unterrichtsverhalten als unbedeutend angesehen werden können, vertraten u. a. auch: B. O. Smith und Meux (1962), Hughes (1959), Flanders (1965) und Bellack u. a. (1966).

Einige Wissenschaftler trafen besondere Vorkehrungen, um die Wirkungen zu verringern, die von Beobachtern und von der Aufnahmeapparatur ausgehen. Zum Beispiel weisen B. O. Smith und Meux (1962) darauf hin, daß die Informationen für die mit ihnen zusammenarbeitenden Lehrer die Versicherung enthielten, daß die Untersuchung keinerlei Bewertung ihrer Lehrfähigkeit bezwecke, daß bei der Veröffentlichung der Forschungsergebnisse völlige Anonymität gewährleistet sei und daß nur Mitglieder des Forscherteams Einblick in die Tonbänder und Manuskripte hätten. Außerdem wurden die Aufnahmegeräte einige Tage vor der Unterrichtsstunde, in der die Beobachtungen stattfanden, im Klassenraum installiert, damit sich Schüler und Lehrer daran gewöhnen konnten.

Ohne die Probleme zu unterschätzen, die durch Beobachter und Aufnahmegerät entstehen, scheint es, daß diese Einflüsse zum großen Teil

ausgeschaltet werden können, wenn entsprechende Vorsichtsmaßnahmen getroffen werden. Da jedoch nicht angenommen werden kann, daß Lehrer und Schüler durch die Anwesenheit eines Beobachters und durch das Vorhandensein eines Aufnahmeapparates trotz derartiger Maßnahmen völlig unbeeinflusst bleiben, sollte man sich an die einfache, offensichtliche Tatsache erinnern, daß es besser ist, »etwas über das Lehrer-Schüler-Verhalten in Erfahrung zu bringen, während sie unter Beobachtung stehen, als überhaupt keine Kenntnisse über das Lehrer-Schüler-Verhalten gewinnen zu können« (Medley/Mitzel 1963, 248).

Die Zuverlässigkeit beim Kodieren

Die Zuverlässigkeit eines Beobachtungsinstruments für das Unterrichtsgeschehen kann definiert werden als »das Ausmaß, in dem die Messung, unter konstanten Bedingungen wiederholt, konstante Ergebnisse erbringt« (Kaplan 1964, 200). Unter den relevanten Bedingungen bei der Verhaltensmessung während des Unterrichts sind von besonderer Bedeutung die Beobachter, die die Messungen vornehmen, sowie die Stabilität des zu messenden Verhaltens. Daher kann die Zuverlässigkeit von Instrumenten zur Beobachtung von Unterrichtsabläufen geschätzt werden durch den Grad der Übereinstimmung zwischen voneinander unabhängigen Beobachtern (Übereinstimmungskoeffizient) und auf Grund der Stabilität der Verhaltensdimensionen unter Beobachtung (Stabilitätskoeffizient). Obwohl beide Arten der Zuverlässigkeit offensichtlich von Bedeutung sind, wird der Stabilitätskoeffizient bei der Erforschung des Unterrichtsverhaltens gegenwärtig kaum berechnet; größere Aufmerksamkeit wird der Schätzung des Ausmaßes an Übereinstimmung zwischen den Beobachtern geschenkt. Die gebräuchlicheren statistischen Indizes für diese Schätzungen sind der Prozentsatz der Übereinstimmung zwischen den kodierenden Personen und der Korrelationskoeffizient; von einigen Wissenschaftlern wird auch die Varianzanalyse angewandt.

Zur Sicherung einer adäquaten Zuverlässigkeit beim Gebrauch von Beobachtungsskalen haben Wissenschaftler bestimmte Verfahrensweisen entwickelt. Von zentraler Bedeutung ist die sorgfältige Entwicklung der Beobachtungsskala selbst, indem man besonders auf die präzise Definition der Kategorien und der Analyse-Einheit und auf die Formulierung von Kodier-Regeln zur Anleitung der kodierenden Personen achtet. Wahrscheinlich resultieren die größten Schwierigkeiten beim zuverlässigen Kodieren aus der unklaren Definition der zu kodierenden Verhaltenseinheit; ohne eine derartige präzise Definition kann kein hoher Grad an Übereinstimmung zwischen den Beobachtern erzielt werden. Die sorgfältige Schu-

lung der Beobachter ist zweifellos ein wesentliches Mittel zur Sicherung der Zuverlässigkeit; die Erfahrung der Forscher bestätigt die Wichtigkeit intensiver Schulungsprogramme für die beobachtenden und kodierenden Personen und für die Notwendigkeit eines systematischen Vorgehens zur Kontrolle der Beobachtungen und Kodierungen.

Statistische Analyse der Daten

Eine detaillierte Diskussion über die Methoden einer statistischen Analyse der Beobachtungsdaten aus dem Unterrichtsgeschehen geht über den Rahmen dieses Beitrags hinaus. Dennoch ist darauf hinzuweisen, daß neuere Theorien und Techniken es ermöglichen, bestimmte Dimensionen von Unterrichtsprozessen zu erforschen, die zuvor nur schwer, wenn überhaupt, statistisch zu erfassen waren. Biddle und Adams (1967) und Bellack u. a. (1966) haben z. B. die zeitliche Verlaufscharakteristik von Ereignissen im Unterrichtsgeschehen statistisch mit Hilfe einer Markoff-Kette beschrieben. In der letzteren Untersuchung versuchten die Wissenschaftler festzustellen, ob bestimmte pädagogische Verhaltensmuster dahin tendieren, andere pädagogische Verhaltensmuster nach sich zu ziehen. Dies wurde statistisch durch die Markoff-Kette beschrieben, mittels derer es möglich ist, die bedingte Wahrscheinlichkeit des Übergangs von einem Verhaltensmuster auf ein anderes zu bestimmen (Kemeny/Snell 1962). Nimmt man die verschiedenen pädagogischen Verhaltensmuster als Einheiten an, dann wurde die Wahrscheinlichkeit des Übergangs von einem Verhaltensmuster zu einem anderen erforscht, um festzustellen, ob ein pädagogisches Verhaltensmuster die Tendenz hat, das jeweils unmittelbar nachfolgende Verhalten zu beeinflussen.

Die statistische Auswertung von Beobachtungsdaten erfuhr eine wesentliche Erleichterung durch den Einsatz von Computern. Gage (1969) berichtete über wichtige Ergebnisse aus der Arbeit von Allen und Snow am Stanford Center for Research and Development in Teaching, die durch den Einsatz von Computern ermöglicht wurden. Allen und Snow entwickelten eine Taxonomie von Verhaltensweisen im Unterricht, indem sie Computer zum Speichern und zum Reproduzieren von Items verwendeten, die sich auf Lehrer- und Schülerverhalten beziehen. Gage beschreibt die Ziele und das Verfahren dieses Projektes:

»Es wurden mehr als 1000 Items gespeichert; diese können vom Computer in einer endgültigen Fassung reproduziert und gedruckt werden, so daß sie den Beobachtern, Bewertern und Wissenschaftlern für die Inhaltsanalyse zur Verfügung stehen. Die Items, die reproduziert werden sollen, können entsprechend vielen differierenden Dimensionen exakt angegeben werden. Die von den Beobachtern oder

den für die Inhaltsanalyse verantwortlichen Wissenschaftlern gewonnenen Daten, die die Häufigkeit, Intensität oder die Verhaltenskorrelate betreffen, die durch diese Items gekennzeichnet werden, können zusammen mit den Items im Computer gespeichert werden. Auf diese Weise werden die Erfahrungswerte gesammelt, die für die Gültigkeit und Zuverlässigkeit der Items bei verschiedenen Lehrertypen, Beobachtern, Lehrinhalten, Effektivitätskriterien und dergleichen relevant sind. Kurz, das Ideal einer universalen Taxonomie des Unterrichtsverhaltens, die für viele Zwecke nützlich ist, kann durch ein auf dem Computer basierendes System – wie in Stanford – erreicht werden« (Gage, 1969, 1452).

II. Beispiele für Beobachtungssysteme

In diesem Abschnitt werden kurz drei Analyse-Systeme beschrieben, die drei Haupttypen von vorhandenen Beobachtungssystemen charakterisieren: (1) Das auf das affektive Geschehen zielende System der Flandersschen Interaktionsanalyse, das sich mit dem sozialen Klima im Klassenraum befaßt (1965); (2) das System von Bellack u. a. zur Analyse des Sprachverhaltens; es enthält vor allem kognitive Dimensionen des Unterrichts (1966); (3) das System von Oliver und Shaver zur Beschreibung von Erziehungsstilen im Sozialkunde-Unterricht, ein multidimensionales System, das sich auf affektive und kognitive Aspekte der Aktivitäten im Unterricht konzentriert (1966).

Ein System zur Analyse affektiver Dimensionen des Unterrichts

Konzeptueller Rahmen: Mit Flanders' System zur Interaktionsanalyse können vornehmlich affektive und interpersonelle Komponenten der Unterrichtsprozesse erfaßt werden. Flanders behauptet, daß die Art und Weise, in der der Lehrer durch seine verbale Kommunikation das Verhalten von Schülern zu beeinflussen sucht, das wichtigste Kennzeichen der Lehrer-Schüler-Beziehung im Unterricht ist. Seine Bemühungen sind darauf gerichtet, zwei vom Lehrer angewandte kontrastierende Beeinflussungsmethoden zu analysieren:

(a) Direkte Beeinflussung, die »aus den verbalen Äußerungen des Lehrers besteht und die die Handlungsfreiheit einschränken, indem entweder die Aufmerksamkeit auf ein Problem gerichtet wird oder der Lehrer sich auf seine Autorität beruft oder beides zusammen erfolgt.

(b) Indirekte Beeinflussung, die sich in solchen verbalen Äußerungen des Lehrers zeigt, die die Handlungsfreiheit des Schülers durch Ermutigung zur verbalen Teilnahme und Initiative am Unterricht vergrößern« (Flanders 1965, 9).

Kategorien und Methode der Analyse: Das Kategoriensystem von Flanders, das für Situationen im Unterricht entwickelt wurde, in denen Lehrer und Schüler aktiv in verbaler Interaktion stehen, beschreibt verbale Handlungen des Lehrers, die direkten und indirekten Einfluß ausüben. Das Beobachtungsschema besteht aus den folgenden zehn Items:

Flanders Kategorien der Interaktionsanalyse

Lehrer- äußerungen	Indirekte Beeinflussung	1. akzeptiert Gefühle 2. lobt oder ermutigt 3. akzeptiert oder bewertet Ideen der Schüler 4. stellt Fragen
	Direkte Beeinflussung	5. trägt etwas vor 6. gibt Anweisungen 7. kritisiert oder rechtfertigt seine Autorität
Schüleräußerungen		8. Schüleräußerungen – Antwort 9. Schüleräußerungen – Eigeninitiative
		10. Ruhe oder Unruhe

Sieben Kategorien sind zur Klassifikation der Lehreräußerungen, zwei zur Klassifikation der Schüleräußerungen vorgesehen. In die zehnte Kategorie fallen Pausen, kurze Ruheperioden oder Phasen der Unruhe. Von den sieben auf die Lehreräußerungen bezogenen Kategorien repräsentieren die Items 1–4 indirekte, die Items 5–7 direkte Einflußnahme.

Die Daten werden von einem Beobachter während des Unterrichts gewonnen. Am Ende jeder 3-Sekunden-Periode entscheidet der Beobachter, welche der zehn Kategorien die Kommunikationsereignisse dieser Periode am besten beschreibt. Er notiert die Numerierungen der jeweiligen Kategorien, während er gleichzeitig die Kommunikation in der nächsten 3-Sekunden-Periode feststellt. Insgesamt erhebt er also ca. 20 Beobachtungen pro Minute. Seine Notierungen bestehen folglich aus einer von oben nach unten in einer Kolonne geschriebenen Sequenz von Zahlen, so daß die ursprüngliche Sequenz von Ereignissen beibehalten bleibt. Wenn sich eine Veränderung im Unterrichtsablauf ergibt, zieht der Beobachter einen Doppelstrich und vermerkt die Zeit. Handlungsperioden auf diese Weise festzustellen ist somit eine zweite Art der Kategorisierung, die zur Ergänzung des Systems zur Klassifizierung verbaler Aussagen

dient. Flanders nennt fünf Handlungsarten: Einführung von neuem Material, Bewertung von Hausaufgaben oder Testen, Klassendiskussion, Beaufsichtigung und Förderung individueller Unterrichtsarbeit, sonstige übliche Schularbeit.

Nach einer bestimmten Beobachtungsperiode überträgt der Beobachter die Numerierungen, die er in Zeitreihen registriert hat, auf eine 10 x 10-Matrix. Die Numerierungen sind in der Matrix jeweils paarweise registriert. Wenn z. B. die Reihenfolge der Numerierungen auf Grund der Beobachtungen lautet: 10, 6, 7, 5, 4, 8, 1, 4, 8, besteht das erste Paar aus den Numerierungen 10-6, das zweite Paar lautet 6-7, das dritte 7-5 etc. Für das erste Paar, 10-6, wird die Registrierung in die Zelle der Reihe 10, Spalte 6 eingetragen, für das zweite Paar, 6-7, in die Zelle der Reihe 6, Spalte 7. Diese Art der Analyse hilft bis zu einem gewissen Grad, die Sequenz der Ereignisse im Unterrichtsablauf festzuhalten. Die Analyse der Matrixdaten liefert verschiedene Informationen über den Interaktionsprozeß im Unterricht: Der Umfang der Lehrer- und Schüleräußerungen, die Zeitspanne, die in Ruhe oder Unruhe verbracht wurde, das Verhältnis von indirektem und direktem Einfluß, den der Lehrer ausübt (I/D-Quotient), die Sequenzen der verschiedenen Arten von Lehrer- und Schüleräußerungen u. ä.

Zuverlässigkeit: Flanders und seine Mitarbeiter entwickelten systematische Verfahrensweisen zur Schulung der Beobachter, um die Zuverlässigkeit des Kodierens zu sichern. Die Zuverlässigkeit der Beobachtungen, die sich durch die Übereinstimmung der Ergebnisse der einzelnen Beobachter ergibt, wird mit Hilfe des Scott-Koeffizienten geschätzt. In einer von Flanders (1965) berichteten neueren Untersuchung lagen die Reliabilitätskoeffizienten von Scott für geschulte Beobachter ausnahmslos über .86.

Anwendungsbereich: Das System der Interaktionsanalyse wurde häufig bei einer Vielfalt von deskriptiven und experimentellen Untersuchungen angewandt, kürzlich auch bei der Lehrerausbildung und -fortbildung, um den Lehrern eine Rückmeldung über ihr eigenes Verhalten im Unterricht zu ermöglichen.

Ein System zur Analyse kognitiver Dimensionen des Unterrichts

Konzeptueller Rahmen: In ihrer Untersuchung »The Language of the Classroom« (1966) konzipierten Bellack und seine Mitarbeiter den Unterrichtsablauf als eine Art Sprachinteraktion (language game), wobei sie Wittgensteins Auffassung von der Sprache als einem durch Regeln gelenkten verbalen Verhalten übernahmen. Die Grundeinheit zur Beschreibung dieser Interaktionen im Unterricht stellt der pädagogische Impuls

(paedagogical move) dar. Diese Impulse werden entsprechend den pädagogischen Funktionen, die sie in der Unterrichtsdiskussion erfüllen, in vier Hauptkategorien eingeteilt:

Strukturieren: Strukturierende Impulse dienen den pädagogischen Funktionen der Setzung des Kontexts für nachfolgendes Verhalten bei entweder erfolgreich beginnender oder schleppend anlaufender Interaktion zwischen Schülern und Lehrern. Die Lehrer beginnen z. B. häufig eine Unterrichtsstunde mit einem strukturierenden Impuls, mit dem sie die Aufmerksamkeit auf das während dieser Stunde zu diskutierende Thema oder Problem lenken wollen.

Auffordern: Pädagogische Impulse in dieser Kategorie sollen verbale Antworten hervorlocken, um die Schüler, von denen Aufmerksamkeit oder eine physische Reaktion erwartet wird, zu ermutigen. Alle Fragen, Befehle, Anweisungen und Bitten sind zu den Aufforderungen zu zählen.

Antworten: Diese pädagogischen Impulse stehen in einer reziproken Beziehung zu den Aufforderungen und treten nur im Zusammenhang mit ihnen auf. Ihre pädagogische Funktion liegt darin, die durch die Aufforderungen ausgedrückten Erwartungen zu erfüllen; so werden Schülerantworten auf Lehrerfragen als Antworten auf die Aufforderungen klassifiziert.

Reagieren: Diese pädagogischen Impulse werden durch ein strukturierendes, aufforderndes, antwortendes oder der Reaktion vorausgehendes Verhalten veranlaßt, sind aber nicht direkt durch sie verursacht. Aus pädagogischer Sicht dienen diese Verhaltensweisen der Modifikation (durch Klärung, Synthetisierung oder Erweiterung) und/oder der Bewertung (positiv oder negativ) dessen, was vorher gesagt wurde. Reaktionen unterscheiden sich von Antworten folgendermaßen: Während eine Antwort immer direkt durch eine Aufforderung hervorgerufen wird, werden die Reaktionen durch vorausgehende pädagogische Impulse lediglich verursacht; z. B. ist die Bewertung einer Schülerantwort durch den Lehrer als eine Reaktion gekennzeichnet.

Die das verbale Lehrer- und Schülerverhalten beschreibenden Impulse treten im Unterrichtsablauf in bestimmten periodischen Verhaltensmustern oder Kombinationen auf, die als Unterrichtszyklen (teaching cycles) bezeichnet werden. Ein typischer Unterrichtszyklus besteht z. B. aus der Aufforderung eines Lehrers, auf die der Schüler mit einer Antwort reagiert, auf die wiederum eine Lehrerreaktion folgt. Der sequentielle Ablauf der Unterrichtszyklen wird durch eine Markoff-Kette beschrieben.

Die Analyse der sprachlichen Interaktionen im Unterricht wäre ohne Beschreibung des Inhalts der pädagogischen Impulse unvollständig. Zwei grundlegende Arten des Inhalts können unterschieden werden:

(a) Inhalte mit fachspezifischer Bedeutung, die sich auf den zur Diskussion stehenden Lehrstoff beziehen (in dieser Untersuchung: Internationaler Handel);

(b) Inhalte mit unterrichtsspezifischer Bedeutung, die sich auf die zugewiesene Aufgabe, auf die Materialien und Verfahrensweisen im Unterrichtsgeschehen beziehen.

Die Inhalte mit fach- und unterrichtsspezifischer Bedeutung werden zusammen mit ihren assoziierten logischen Bedeutungen beobachtet und registriert, die sich auf die mit den fachspezifischen und unterrichtsspezifischen Inhalten – wie definieren, interpretieren, Fakten darlegen, erklären, Meinungen äußern und begründen – in Verbindung stehenden kognitiven Prozesse beziehen. Wenn ein Schüler also die Fragen eines Lehrers beantwortet, indem er eine Definition von »Preis« gibt, ist der Impuls des Schülers als eine *Antwort* zu kodieren, die fachspezifische Bedeutung als *Preis*, und der logische Prozeß als *definieren*. Zusätzlich wird der Schüler als der Sprecher gekennzeichnet und die Länge seiner Antwort nach der Anzahl der Schreibmaschinenzeilen gemessen.

Kodierungsverfahren: Das Kodieren erfolgt nach Gesichtspunkten des Beobachters, der den pädagogischen Bedeutungsgehalt dem verbalen Verhalten des Redners entnimmt. Die mit der Kodierung beauftragten Personen hören die Bandaufnahmen ab, die vom Unterrichtsgeschehen aufgenommen wurden, und verwerten ebenso die Niederschriften der Tonbandaufnahmen. Jeder pädagogische Impuls wird folgenden Analysekatégorien entsprechend kodiert:

- (1) Sprecher
- (2) Art der pädagogischen Impulse
- (3) fachspezifische Bedeutung
- (4) fachspezifisch-logische Bedeutung
- (5) Antwortenlänge bei (3) und (4)
- (6) unterrichtsspezifische Bedeutung
- (7) unterrichtsspezifisch-logische Bedeutung
- (8) Antwortenlänge bei (6) und (7)

Das folgende Beispiel stellt einen kodierten pädagogischen Impuls dar:

L / STR / IE / ERK / 4 / V / F / 2
 (1) (2) (3) (4) (5) (6) (7) (8)

Diese Angaben werden wie folgt interpretiert: Ein *Lehrer* (1) gibt einen *strukturierenden* (2) Hinweis, indem er etwas über *Import* und *Export* (3) in einer Länge von *vier* (5) Schreibmaschinenzeilen *erklärt* (4); ferner informiert er die Klasse *faktisch* (7) darüber, wie sie im weiteren Verlauf des Unterrichtsgeschehens *verfahren* (6) soll, und zwar in einer Länge von *zwei* (8) Zeilen. Im folgenden wird eine kurze Zusammenfas-

sung der Analysekategorien gegeben, wobei die acht Numerierungen mit den oben genannten übereinstimmen:

1. Sprecher
 - Lehrer (L)
 - Schüler (S)
 - audiovisuelles Gerät (A)
2. Art des pädagogischen Impulses
 - strukturieren (STR)
 - auffordern (AUF)
 - antworten (ANT)
 - reagieren (REA)
3. Fachspezifische Bedeutungen (Lehrinhalt – in dieser Untersuchung: Internationaler Handel; die Kategorien basieren auf einer Inhaltsanalyse eines Textes über den internationalen Handel, der in den an der Untersuchung teilnehmenden Klassen zugrunde gelegt wurde)
 - Handel (HA)
 - Produktions- und/oder Spezialisierungsfaktoren (PSF)
 - Importe und/oder Exporte (IE)
 - ausländische Investitionen (AI)
 - Handelsbarrieren (HB)
 - Förderung des freien Handels (FFH)
 - Kriterien, die für den Handel von Bedeutung sind (KHB)
 - Gründe, die gegen den Handel sprechen (GH)
4. Fachspezifisch-logische Bedeutungen (kognitive Prozesse, die bei der Behandlung von Lehrinhalten miteingeschlossen sind)
 - Analytische Prozesse
 - definieren (DEF)
 - interpretieren (INT)
 - Empirische Prozesse
 - formulieren von Faktoren (F)
 - erklären (ERK)
 - Evaluative Prozesse
 - Meinungen äußern (MÄ)
 - begründen (BEG)
5. Antwortenlänge in Schreibmaschinenzeilen bei 3. und 4.
6. Unterrichtsspezifische Bedeutungen (Faktoren, die auf die Klassenführung bezogen sind)
 - Zuweisung einer Aufgabe (ZA)
 - Material (M)
 - Person (P)
 - Verfahren (V)

Aussage (A)
 logischer Prozeß (LP)
 Sprachtechnik (ST)
 allgemeine Handlungen (AH)
 verbale Handlungen (VH)
 physische Handlungen (PH)
 kognitive Handlungen (KH)
 emotionale Handlungen (EH)

7. Unterrichtsspezifisch-logische Bedeutungen (kognitive Prozesse, die mit unterrichtsspezifischen Bedeutungen assoziiert sind)
 analytische, empirische, evaluative Prozesse (wie bei 4.)
 Einschätzen (Bezug zur Metakommunikation, gewöhnlich eine evaluative Reaktion)
 positiv (POS)
 zustimmend (ZU)
 wiederholend (WDH)
 genau beschreibend (GB)
 nicht zustimmend (NZU)
 negativ (NEG)
 außer-logische Prozesse
 ausführen (AUS)
 anleiten (ANL)
8. Antwortenlänge in Schreibmaschinenzeilen bei 6. und 7.

Zuverlässigkeit: Um die Zuverlässigkeit des Kodierens zu schätzen, kodierten zwei Gruppen des Forscherteams Stichproben von Protokollen, die nach dem Zufallsprinzip ausgewählt worden waren, und verglichen die Ergebnisse miteinander. Der Prozentsatz der Übereinstimmung wurde für jede der fundamentalen Kategorien des Analyse-Systems berechnet, und zwar hinsichtlich der Anzahl und Dauer der pädagogischen Impulse. Das Ergebnis erbrachte einen übereinstimmend hohen Zuverlässigkeitsgrad für alle Kategorien. Die Übereinstimmung lag zwischen .84 und .96.

Anwendungsbereich: Mit geeigneten Modifikationen in den Kategorien, die als »fachspezifische Bedeutung« klassifiziert sind, liefert dieses Verfahren eine nützliche Technik zur Erforschung des Unterrichtsverhaltens in einer Vielzahl von Fächern der Primar- und Sekundarstufe. Vor kurzem wurden am »Teachers College« Dissertationen beendet, in denen diese Methode der Analyse zur Erforschung des Mathematikunterrichts in der Sekundarstufe und des Leseunterrichts in der Primarstufe verwendet wurde.

*Ein System zur Analyse affektiver und kognitiver Dimensionen
des Unterrichts*

Konzeptueller Rahmen: Im Zusammenhang mit ihrer Untersuchung »Teaching Public Issues in the High School« (1966) entwickelten Oliver und Shaver ein Beobachtungssystem zur Analyse von zwei Unterrichtsstilen: Lehrervortrag und sokratisches Gespräch. Der theoretische Rahmen, in dem diese beiden Unterrichtsweisen definiert sind, beinhaltet drei Dimensionen der Interaktion:

- a) kognitive Dimension,
- b) affektive oder sozial-emotionale Dimension,
- c) Verfahrensdimension.

Was die kognitiven Aspekte betrifft, »tendiert der Lehrervortrag zur Darstellung. Hierbei wird angenommen, daß der Tatbestand eindeutig dargestellt werden kann und daß der Lehrer nur Informationen vorzutragen und zu erklären hat oder eine analytische Struktur vorgibt, durch die die Informationen gegliedert werden können« (Oliver/Shaver 1966, 178). Im Gegensatz dazu »ist der sokratische Stil zweifellos dialektisch. Er setzt voraus, daß das Problem nur in einem entgegengesetzte Aspekte beinhaltenden Kontext geklärt werden kann, in dem verschiedene Standpunkte dargestellt und verteidigt werden« (a. a. O., 178). Hinsichtlich der affektiven Aspekte der Interaktion ist es naheliegend, daß sowohl der sokratisch arbeitende als auch der vortragende Lehrer den Schüler unterstützen muß. Da jedoch »die offene Kontroverse auf dem kognitiven Niveau in den affektiven Bereich übergeht, tendiert die sokratische Methode dahin, daß eine Diskussion stark mit negativen Affekten aufgeladen wird« (a. a. O., 179). Da der Lehrer jedoch im Zentrum des Dialogs mit den Schülern steht, werden keine systematischen Unterschiede zwischen vortragenden und sokratisch unterrichtenden Lehrern in dieser Dimension erwartet.

Kategorien für die Analyse: Oliver und Shaver entwickelten eine Reihe von Kategorien, die für die oben beschriebenen Unterrichtsstile des Lehrervortrags und des sokratischen Gesprächs bedeutsam sind. Ihr Untersuchungsinstrument besteht aus den folgenden drei Gruppen von Kategorien:

Affektive und sozial-emotionale Kategorien

1. Solidarität
2. geringer positiver Affekt
3. Entspannung
4. Spannung
5. geringer negativer Affekt

6. Widerspruch
neutral (keine affektive Mitteilung)

Kognitive Kategorien

7. Hinweis auf Inkonsistenz
8. Beschreibung
9. Evaluation
10. Wiederholung, Zusammenfassung, Herausstellung des Wesentlichen
11. Klärung
12. Hinweis auf Analogien
nicht-kognitiv

Verfahrenskategorien

13. Anordnung aufgabenorientierten Verhaltens
14. Kontrolle abweichenden Verhaltens

Der Beobachter, der dieses Instrument benutzt, muß auf die kognitive oder sonstige für den Unterrichtsverlauf relevante Bedeutung in jeder Aussage oder in jeder Gedankeneinheit schließen, die vom Lehrer zum Ausdruck gebracht wird. Jede derartige einstuftbare Handlung wird in einer kognitiven oder Verfahrenskategorie und in einer affektiven Kategorie eingetragen. Die Funktionen der kognitiven Kategorien werden folgendermaßen beschrieben: »Die primäre Funktion der kognitiven Kategorien (7–12) liegt darin, Fragen über Unterschiede in bezug auf den intellektuellen oder logischen Inhalt in den nach den beiden verschiedenen Stilen stattfindenden Unterrichtsabläufen zu beantworten. Von zentraler Bedeutung für die Unterscheidung zwischen den beiden Stilen ist das Ausmaß, in dem sich die Lehrer mit *deskriptiver Information* bei kontroversen Fragen im Verhältnis zu *Werturteilen* befassen, die über diese Fragen abgegeben werden. Die Kategorien 8 und 9 sind speziell dafür vorgesehen, Unterschiede dieser Art aufzudecken. Die Kategorie 7 soll die Versuche des Lehrers zum Ausdruck bringen, bei den Schülern persönliche Wertkonflikte dadurch hervorzurufen, daß er den Eindruck erweckt, gegensätzliche Urteile in gleichen Situationen abzugeben (das kontroverse Problem und eine Analogie). Die Kategorie 12 (Analogie) hat so eine offensichtliche Bedeutung. Die Kategorien 10 und 11 wurden miteinbezogen, um das kognitive Subsystem zu vervollständigen« (Oliver/Shaver 1966, 291–292).

Zuverlässigkeit: Die Beobachter wurden anhand von Tonbandaufnahmen der Klassendiskussionen sorgfältig in der Anwendung des Verfahrens geschult. Eine Bearbeitung des Chi-Quadrat-Verfahrens wurde zur Schätzung der Übereinstimmung zwischen den Beobachtern verwendet. Alle Chi-Quadrat-Werte lagen deutlich unter dem 5 %-Niveau, das die Wissenschaftler als Kriterium für die Untersuchung angenommen hatten.

Anwendungsbereich: Obwohl dieses analytische Verfahren zur Verwendung im Fach Sozialkunde entwickelt wurde, in dem öffentliche Probleme im Zentrum der Diskussion stehen², nehmen die Autoren an, daß es ebenso zur Analyse des Unterrichtsgeschehens in anderen Fächern, z. B. im naturwissenschaftlichen und Literaturunterricht, verwendet werden kann.

III. Probleme und Konsequenzen in der Entwicklung und im Gebrauch von Systemen zur Analyse des Lehrer- und Schülerverhaltens

In diesem Abschnitt wendet sich die Diskussion zwei Problemen und Konsequenzen zu, die den Erziehern bei der Entwicklung und Anwendung von Beobachtungssystemen im Unterricht begeben.

Beziehungen zwischen den Beobachtungssystemen

Das erste Problem erwächst aus dem Bestreben der Forscher, ein neues Analyse-System für nahezu jede neue Untersuchung zu entwickeln, wohl in der Annahme, die von anderen entwickelten Verfahren seien ungeeignet oder inadäquat. Es mag in der Tat zutreffen, daß Aspekte untersucht werden sollen, für die eine Beobachtungsskala noch nicht entwickelt wurde, oder daß die vorhandenen Systeme nur eine geringe Gültigkeit und Zuverlässigkeit besitzen. Die Entwicklung zahlreicher verschiedener Beobachtungssysteme hat zur Folge, daß den Wissenschaftlern nur sehr wenige Daten zur Verfügung stehen, die einen Vergleich verschiedener Untersuchungen zulassen.

Die umfangreiche Konstruktion von Beobachtungssystemen ist zweifellos das Ergebnis unterschiedlicher Konzeptionen der Forscher. In diesem Zusammenhang sei daran erinnert, daß »in den frühen Stadien der Entwicklung jeder Wissenschaft verschiedene Menschen, die mit dem gleichen Bereich von Phänomenen, nicht aber mit denselben Einzelphänomenen konfrontiert sind, diese unterschiedlich beschreiben und interpretieren« (Kuhn 1962, 17). So kann von den Wissenschaftlern auf dem Gebiet der Unterrichtsforschung – einem Sektor, der sich gerade als wissenschaftlicher Forschungsbereich abzuzeichnen beginnt – nicht erwartet werden, daß sie ihren Forschungen ein gemeinsames theoretisches Konzept zugrunde legen. In diesem Stadium der Entwicklung sind in Widerspruch stehende und konkurrierende Ansätze zur Beschreibung von Unterrichtsabläufen zu erwarten und sogar zu unterstützen.

Jedoch schließt dieser Stand der Dinge nicht die Möglichkeit von Un-

tersuchungen zur Bestimmung von Faktoren oder Dimensionen aus, die den verschiedenen Beobachtungssystemen zugrunde liegen, so daß Differenzen und Ähnlichkeiten zwischen ihnen klarer unterschieden werden können. Ein faktorenanalytisches Verfahren, das dem Vergleich von Beobachtungssystemen zum Zweck der Feststellung von Unterschieden zwischen den ihnen zugrunde liegenden Dimensionen dient, wurde von Gage vorgeschlagen (1969, 1450):

»Die Verhaltensweisen von Lehrern in einer großen Stichprobe können aufgrund vieler Variablen solcher Art, wie sie von Flanders (1965), Smith u. a. (1962), Bellack u. a. (1966), Medley und Mitzel (1958) u. a. entwickelt wurden, erfaßt werden. Dann können die Interkorrelationen der aus den Variablen gewonnenen Punktwerte der Faktorenanalyse unterzogen werden. Die ermittelten Faktoren können die Dimensionen in relativ knapper oder einfach strukturierter Form definieren.«

Wenn vergleichende empirische Untersuchungen von vorhandenen Systemen verfügbar werden, wird es möglich sein, viele Aspekte berücksichtigende Beobachtungsskalen zu konstruieren, die Konzepte einer Vielfalt verschiedener Systeme in sich vereinen. Meux (1967) hat z. B. die Entwicklung eines neuen multidimensionalen Systems vorgeschlagen, um Einheiten mit lernzielorientierten Inhalten und innerhalb dieser Einheiten Perioden, in denen initiiert und reagiert wird, sowie fachspezifische und unterrichtsspezifische Aspekte, Gruppenprozesse und Denkprozesse einzubeziehen. Er ist der Auffassung, daß Unterrichtsstrategien konstruiert werden können, die weit effektiver als die besten der zur Zeit von unseren Lehrern verwendeten sind, wenn Kategorien wie die oben genannten von verschiedenen Beobachtungssystemen in dem vorgeschlagenen multidimensionalen System kombiniert werden. »Diese Möglichkeit neuer Kombinationen von gegenwärtigen Praktiken ist von einigen Autoren übersehen worden, die der Auffassung sind, daß Beobachtungsverfahren uns prinzipiell niemals über die gegenwärtige Praxis hinausbringen werden« (Meux 1967, 549–550).

Mit ähnlichen Zielsetzungen hat Biddle (1967) eine andere Variante eines Idealsystems entworfen, das drei umfassende Konzepte miteinbeziehen soll:

- (a) Übergreifende Konzepte für das Verhalten im Unterricht, wobei strukturelle und funktionelle Aspekte des sozialen Systems – dargestellt durch den Klassenverband – mitberücksichtigt werden sollen;
- (b) Konzepte für die beobachtbaren, sich gegenseitig beeinflussenden Handlungen der Unterrichtsteilnehmer;
- (c) Konzepte für den Sprachgebrauch in der Klasse einschließlich Syntax und Phonologie.

Ob in diesem Entwicklungsstadium der Unterrichtsforschung die Wissenschaftler ihr Augenmerk mehr auf die Konstruktion von multidimensionalen Verfahren wie die von Meux und Biddle vorgeschlagenen richten sollen, oder ob das primäre Bemühen auf die Konstruktion von Beobachtungssystemen von begrenztem Rahmen konzentriert werden soll, ist eine umstrittene Frage. Es ist jedoch wenig sinnvoll, wenn Forscher sich auf das eine oder andere Vorgehen beschränken. In Anbetracht unseres begrenzten Wissens über den Unterrichtsprozeß könnte man vernünftigerweise erwarten, daß beide Ansätze zu einem besseren Verständnis der Unterrichtsprozesse beitragen.

Mögliche Anwendungsbereiche für Beobachtungsverfahren

Die Beobachtung von Lehrer- und Schülerverhalten ist zu einer anerkannten Methode für die Sammlung von Daten in der Unterrichtsforschung geworden. Die mit Hilfe der Beobachtungstechniken gewonnenen Daten geben Anlaß zu der Hoffnung, daß sich unser Verständnis für das komplexe Leben im Klassenraum vertieft. Der zunehmende Gebrauch von Beobachtungstechniken in der Unterrichtsforschung hat Pädagogen, die in der Lehrerbildung und Schulaufsicht tätig sind, zu dem Vorschlag angeregt, die gleichen Beobachtungsverfahren für die Lehrerausbildung und -fortbildung einzusetzen. Dieser Vorschlag wirft mehrere Probleme auf, die wenigstens kurz erörtert werden sollen.

Für den Einsatz von Beobachtungssystemen in der Lehrerausbildung und -fortbildung wird von Simon und Boyer (1967) folgende Begründung gegeben:

»Zunächst stellen die Systeme einen Spiegel für den Lehrer dar, um eine Rückmeldung über sein eigenes Lehrverhalten in den im angewandten System berücksichtigten Dimensionen zu erhalten. Dieses Feedback stellt für die Lehrer eine günstige Gelegenheit dar, ihr Verhalten aufgrund der Informationen darüber, wie sie sich im Unterricht verhalten, selbst zu modifizieren. Ferner – und dies ist eventuell ein noch wichtigerer Aspekt – sind viele dieser Systeme aufgrund einer theoretischen Dimension konstruiert worden, die Verhaltensweisen beinhaltet, die bei der Verwendung im Unterricht eine Hilfe zur Förderung der Entwicklung der Schüler darstellen sollen, wie es zur Zeit in den Schulen in der Weise noch nicht der Fall ist. Wenn ein Lehrer eines dieser Systeme anwendet, erhält er ein Feedback über die Verhaltensweisen, die er *nicht* zeigt, wie auch über die, die er zeigt. Dem Lehrer bietet sich somit eine Möglichkeit, neue Verhaltensweisen zu lernen und auf diese Weise sein Verhaltensrepertoire mit Hilfe von Methoden zu erweitern, die den Lehrern im allgemeinen nicht zur Verfügung stehen« (Simon/Boyer 1967, 19–20).

Kaum einer würde die Wichtigkeit konzeptueller Instrumente für die

Ausbildung und Fortbildung von Lehrern leugnen, die sie befähigen, den Unterrichtsprozeß zu analysieren und auf diese Weise das Verständnis für ihre Aufgaben als Lehrer zu vertiefen. Wenn die verschiedenen Analyseverfahren, die im Zusammenhang mit Forschungsprojekten entwickelt wurden, den Lehrern als Instrumente vorgestellt werden, mit denen sie den Unterricht von verschiedenen Perspektiven her betrachten können, sind die Lehrer in der Lage, aus den Analysen große Vorteile zu ziehen. In einem kürzlich erschienenen Bericht der American Association of Colleges for Teacher Education (1968) wird der Vorschlag gemacht, eine Phase in der Lehrerausbildung einer solchen Analyse des Unterrichts zu widmen.

Jedoch ist ein Wort der Vorsicht angebracht. Die Versuchung, Beschreibungen des Lehrerverhaltens in Vorschriften umzuwandeln, die von Lehrern zu befolgen sind, muß vermieden werden. Denn bisher konnten nur wenige Korrelationen zwischen den Leistungen der Schüler und vielen Unterrichtsvariablen in den vorhandenen Beobachtungssystemen festgestellt werden. Obwohl die Beobachter, die die verschiedenen Analyseverfahren anwenden, dem Lehrer zuverlässige Beschreibungen vieler schwieriger Dimensionen der Unterrichtsabläufe liefern können, dürfen sie selbstverständlich nicht vorschreiben, wie Lehrer auf der Basis dieser Beschreibungen unterrichten sollen. Denn Entscheidungen hinsichtlich des sich daraus ergebenden Unterrichtsverhaltens sind vom Wissen um Konsequenzen verschiedener Verfahren und vom Urteil hinsichtlich des Wertes dieser Konsequenzen abhängig.

Abgedruckt auch in: M. C. Wittrock/D. E. Wiley, *The evaluation of instruction, issues and problems*, New York: Holt, Rinehart and Winston 1970, 221–238.

1 Sehr zum Leidwesen vieler unwilliger Beamter der Schulverwaltung gestehen wir jedoch daß dies ein tunlicher Ausgangspunkt wäre.

2 Manches spricht dafür, daß dies ein vernünftiges Verfahren ist. Vgl. J. S. Bekker 1962; Miller: *Income and Higher Education*, in: S. J. Muskin 1962; T. Schultz 1961.

GENE V. GLASS: Die Entwicklung einer Methodologie der Evaluation

Übersetzung von Hannes Graudenz (Dipl.-Psych.) und dem Herausgeber.

Originaltitel: *The growth of evaluation methodology*

Der deutschen Übersetzung liegt das dem Herausgeber vom Autor zugesandte Manuskript zugrunde.

1 Ein »allgemeines Phänomen« ist nachgewiesen oder kann entdeckt werden in einem weiten Feld von scheinbar verschiedenen Erscheinungen und wird als Kriterium zur Prüfung eines wissenschaftlichen Begriffs herangezogen. Ohne eine solche Qualifikation würde es bereits »Einschätzung wissenschaftlicher Wahrheit« bedeuten, empirisch festzustellen, daß man seine Schlüssel verloren hat. Der Begriff der Generalisierbarkeit von erwarteten Ergebnissen ist wichtig für die Unterscheidung von Evaluation und Forschung; er ist auch von großer praktischer Bedeutung beim Entwurf einer Evaluations-Untersuchung (vgl. Stake 1969).

2 In diesem Abschnitt beziehe ich mich weitgehend auf die Geschichte der North Central Association von Calvin O. Davis (1945).

3 Die Pionierarbeit von Joseph M. Rice mag manchem zu dieser Zeit bekannt gewesen sein, wurde aber wahrscheinlich eher als tendenzieller Journalismus denn als pädagogische Forschung angesehen.

ARNO A. BELLACK:

Methoden zur Beobachtung des Unterrichtsverhaltens von Lehrern und Schülern

Übersetzung von Dorothea Szymanski (Dipl.-Psych.).

Originaltitel: *Methods for observing classroom behaviour of teachers and students*, in: K. Ingenkamp: *Methods for the evaluation of comprehensive schools*, Weinheim, Berlin, Basel 1969, 187–215.

1* Zum gegenwärtigen Zeitpunkt liegen bereits 16 Bände mit Instrumenten zur Unterrichtsbeobachtung (79 Instrumente) vor (vgl. Bibliographie).

2* Vgl. dazu auch Wulf 1972a.