

Arnold, Karl-Heinz; Vollstädt, Witlof

Arbeits- und Sozialverhalten in der Schule. Möglichkeiten und Grenzen ihrer Beurteilung durch "Kopfnoten"

Die Deutsche Schule 93 (2001) 2, S. 199-209



Quellenangabe/ Reference:

Arnold, Karl-Heinz; Vollstädt, Witlof: Arbeits- und Sozialverhalten in der Schule. Möglichkeiten und Grenzen ihrer Beurteilung durch "Kopfnoten" - In: Die Deutsche Schule 93 (2001) 2, S. 199-209 - URN: urn:nbn:de:0111-pedocs-275518 - DOI: 10.25656/01:27551

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-275518>

<https://doi.org/10.25656/01:27551>

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, vertreiben 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

Karl-Heinz Arnold und Witlof Vollstädt

Arbeits- und Sozialverhalten in der Schule

Möglichkeiten und Grenzen ihrer Beurteilung durch
„Kopfnoten“

Die in unserem Schulsystem erreichten Lernleistungen sind wieder einmal in die öffentliche Kritik geraten. Mittelmäßige Ergebnisse der bundesdeutschen Schülerinnen und Schüler bei internationalen Vergleichstests wie z. B. bei der TIMS-Studie (Third International Mathematics and Science Study) haben eine engagierte bildungspolitische und schulpädagogische Diskussion ausgelöst. Dabei führte die Suche nach Erklärungen für das unerwartet schwache Abschneiden der Bundesrepublik zu einem Thema, dem über viele Jahre im erziehungswissenschaftlichen Diskurs wenig Aufmerksamkeit gewidmet worden war: dem Zusammenhang von Leistungsermittlung, -bewertung und -förderung. Die Problematik der Beurteilung schulischer Leistungen durch Zensuren und Notenzeugnissen galt offenbar seit Ingenkamp (1974) als hinreichend genau erforscht. Dazu steht allerdings die relative Folgenlosigkeit seiner Kritik im deutlichen Widerspruch. Obwohl er eindrucksvoll nachgewiesen hatte, dass Zensuren die Selektionsfunktion im Rahmen des staatlichen Berechtigungswesens nur sehr unzulänglich erfüllen, weil sie nicht objektiv, reliabel und valide genug sind und auch nicht sein können, blieb die Praxis der schulischen Leistungsbeurteilung und -rückmeldung – zumindest im Sekundar-schulbereich – davon weitgehend unberührt.

Nunmehr hat erneut eine Diskussion eingesetzt, ob und welche Testverfahren und standardisierten Methoden zur Leistungsmessung geeignet sind, um die Objektivität und Validität der schulischen Bewertungen zu verbessern. Die Orientierung an internationalen Schulleistungsstudien sowie an landeseinheitlichen Testverfahren bzw. Vergleichsarbeiten bietet jedoch nur eine Teillösung: schulübergreifende Bewertungsstandards erhalten ein höheres Gewicht. Aus pädagogischer Sicht bleibt abzuwarten, wohin diese Suche führen wird.

Zumindest sollte mit gleicher Intensität darüber nachgedacht werden, welche Ziele mit schulischem Lernen heute und künftig verfolgt und welche Qualifikationen und Kompetenzen dabei erworben werden sollen. Diese Sicht führt zwangsläufig zur Vermutung, dass neben fachlichen und fachdidaktischen Qualitätsmerkmalen des Unterrichts auch allgemeine pädagogische Merkmale des schulischen Lernens für das Erreichen höherer Fachleistungen bedeutsam sind und verstärkter Aufmerksamkeit bedürfen. Dies könnte wichtiger sein als alle Versuche, die Zensurierung von Leistungen und Verhalten in der Schule aufzuwerten, zumal nach wie vor unter schulischer Leistung recht Unterschiedliches verstanden wird (vgl. Tillmann/Vollstädt 1999). Trotzdem: Der Ruf nach einer „strengerer“ Zensurierung ist unüberhörbar. Die Hoffnung auf die leistungsstimulierende Kraft und den Informationswert von Zensuren scheint so stark zu

sein, dass mit großer Presseresonanz über die Wiedereinführung der „Kopfnoten“ nachgedacht wird, wie dies in Sachsen bereits geschehen ist.

Lehmann/Ziegenspeck (2000) stellen Kopfnoten in die Nähe zu sog. Schlüsselqualifikationen. Dies ist nicht unproblematisch; es lässt sich eine Reihe diagnostischer und pädagogischer Gründe anführen, z. B. Mitarbeit im Unterricht als „Regelorientierung“ im Setting der Schule aufzufassen, nicht jedoch als Kompetenzindikator (vgl. Arnold/Jürgens 2001, S. 98). Selbst ein Blick zurück in die Geschichte der Kopfnoten verweist auf deren begrenzte Aussagekraft, wenn es um Kompetenzen geht.

1. „Kopfnoten“ in der Geschichte der Zeugnisgebung

In der Debatte über Kopfnoten wird leicht übersehen, dass in der historischen Entwicklung der Schulzeugnisse die *allgemeine Beurteilung des Schülers* im Vordergrund stand: „Es [das Zeugnis] war bis in die Neuzeit hinein nichts weiter als eine Bescheinigung über Schulbesuch und Fleiß für Stipendiaten.“ (Ingenkamp/Laux 1990, S. 57) Erst die spätere Ausdifferenzierung der pädagogischen Diagnostik sowie die Akzentuierung der regelmäßigen Berichtsgegenüber der Zertifizierungsfunktion ließ die fachleistungsbezogenen Beurteilungen zum zentralen Aussagebereich der Zeugnisse werden. So zitiert Dohse (1967, S. 34) aus den in Wismar gegen Ende des 18. Jahrhunderts geltenden Schülergesetzen, dass „das Betragen eines jeden Schülers in Ansehung seiner Ordnungsliebe, seines Fleißes, seiner wohlanständigen Aufführung und seines Gehorsams“ von den Lehrern regelmäßig zu protokollieren und zu beurteilen sei.

Gegen Ende des 19. Jahrhunderts entwickeln sich aus diesen freien Beschreibungen *skalenorientierte Kennzeichnungen* für die allgemeine Beurteilung in Zeugnissen, die nun auch Zensuren für die Fachleistungen enthalten. Zumeist werden „Betragen“ und „Fleiß“ als Beurteilungsgegenstände gewählt (vgl. Dohse 1967, S. 53) und mit vorgegebenen Attributen eingeschätzt, die vor allem im unteren Merkmalsbereich Abweichungen zu den verbalen Zensurenstufen aufweisen. Hier deuten sie zumeist emotionale oder pädagogische Reaktionen (also keine neutralen Bewertungen) wie „unbefriedigend“ oder „tadelnswert“ an. Da es sich hier um „Allgemeine Beurteilungen“ handelte, wurden diese von den „speziellen“, d.h. von den fachleistungsbezogenen abgegrenzt und im Anfangsteil des Zeugnisses notiert – gewissermaßen in der „Kopfzeile“ des Zeugnisblattes. Daher resultiert die Bezeichnung „Kopfnoten“.

Vermutlich führte vor allem die zunehmend stärkere Fachorientierung der höheren Schulbildung zu einem Bedeutungsverlust der allgemeinen pädagogischen Beurteilungsfunktionen. So verweisen Ingenkamp/Laux (1990, S. 57) und Rauschenberger (1999, S. 67) auf die Abschaffung der sog. Führungsnoten im Gymnasium im Jahr 1930, der mit dezidiert pädagogischen Argumenten entgegengetreten wurde: „Aber es ist ein verhängnisvoller Irrweg, das Zeugnis zu einem bloßen Dokument des intellektuellen Wissens zu machen“ (Wolff 1932; zit. nach Ingenkamp/Laux 1990, S. 57/58).

Aus der pädagogisch-diagnostischen Forschung der 20er Jahre lagen bereits deutsche Untersuchungen zur methodischen Problematik der schulischen Leistungsbeurteilung vor – z.B. die Mannheimer Aufsatzstudie von Lämmer-

mann aus dem Jahr 1927 oder die auch heute noch vielzitierte Diktatstudie von Zillig aus dem Jahr 1928. In letzterer wurde unter anderem nachgewiesen, dass Lehrer in Diktaten leistungsstarker Schüler mehr Fehler übersehen als in jenen der schwachen Schüler, was ein beträchtliches Fairnessproblem darstellt (vgl. Ingenkamp/Laux 1990, S. 89 u. 92). Unter dem Stichwort „Beobachtungsbogen oder freie Charakteristik“ sind in den drei Anfangsjahrzehnten des 20. Jahrhunderts eine beachtliche Vielzahl von Instrumenten entwickelt worden, die eine strukturierte und zum Teil verhaltensnahe Einschätzung der unterrichtlichen Lernaktivitäten, lernrelevanter Einstellungsmerkmale und allgemeiner Fähigkeiten ermöglichen sollten (vgl. Ingenkamp/Laux 1990, S. 269-310). Erste Einsichten in Beurteilungstendenzen und Urteilsverzerrungen wurden damals bereits formuliert. Gleichwohl dominierte in der pädagogischen Diskussion der Wunsch, die Individualität der Schüler möglichst differenziert darzustellen, was man durch freie Schülercharakterisierungen meinte eher erreichen zu können.

2. Kopfnoten in der Bundesrepublik und in der DDR

Nach dem zweiten Weltkrieg wurden sowohl in den Ländern der BRD als auch in der DDR die Kopfnoten als Zeugniselemente wieder eingeführt. Als Beurteilungskategorien fungierten die Bezeichnungen „Betragen“, „Fleiß“, „Aufmerksamkeit“ und „Ordnung“, teilweise ergänzt um oder ersetzt durch „Beteiligung am Unterricht“. In den Sechziger- und Siebzigerjahren wurden in den meisten Ländern der Bundesrepublik – nicht jedoch in der DDR – die Kopfnoten abgeschafft, da die curricularen Regelungen weder systematische Lerngelegenheiten für diese Verhaltensbereiche enthielten noch eine empirisch hinreichend sichere Handhabung dieser Beurteilungsaufgaben gewährleistet werden konnte. Die moderne sozialpsychologische Forschung hatte eine Fülle von Untersuchungsergebnissen bereit gestellt, die problematische Nebenwirkungen und Fehlerquellen in der Beurteilung allgemeiner Person- und Verhaltensmerkmale aufzeigten.

In der DDR stand die Bewertung des Verhaltens mit Zensuren offiziell nie ernsthaft zur Debatte. Die Anweisung des Ministeriums für Volksbildung vom 10.11.1961 hatte die Vergabe von Zensuren für das Gesamtverhalten, für Betragen, Fleiß und Mitarbeit sowie Ordnung einheitlich für alle Jahrgänge, allerdings nicht für das Reifezeugnis, festgelegt. Grundlage für die Zensurierung waren die Forderungen der Schulordnung an die Schüler. Fachzensuren sollten ausschließlich für die nachgewiesenen Leistungen erteilt werden, für die Einschätzung des Verhaltens gab es Kopfnoten (vgl. Döbert/Geißler 2000, S. 118f.). Diskussionen in der Praxis und Widerstände in der Wissenschaft waren trotzdem nicht völlig zu verhindern. Um ihnen ein Ventil zu verschaffen, wurden ab Schuljahr 1975/76 in den Klassen 11 und 12 der erweiterten Oberschulen (gymnasiale Oberstufe) sämtliche Kopfnoten und ab 1. September 1978 auch in den Klassen 1 bis 10 wenigstens die Noten für das Gesamtverhalten abgeschafft. Betragen, Fleiß, Mitarbeit und Ordnung mussten dort allerdings weiterhin mit Noten bewertet werden.

Die in den Zensierungsanweisungen (vgl. Ministerium für Volksbildung der DDR 1980; siehe auch Witzlack 1986) verwendete Terminologie zeigte ein diagnostisch aufschlussreiches Oszillieren: so wurde explizit die „Verhaltens-

zensierung“ von der „Gesamteinschätzung“ abgegrenzt, beide sollten jedoch von der „Entwicklung der Gesamtpersönlichkeit des Schülers“ ausgehen, wobei die Gesamteinschätzung „entwicklungsfördernd“ wirken sollte, indem sie „die positiven Erscheinungen im Verhalten hervorhebt, kritische Wertungen ausgewogen und pädagogisch taktvoll trifft und konkrete Hinweise zur weiteren Entwicklung des Schülers gibt“. Auch wenn die Parallele zum Widerspruch reizt: ähnliche Konzepte finden sich in den Begründungen für Lernentwicklungsberichte, die typisch „westlicher“ Provenienz sind – allerdings nur, wenn deren reformpädagogische Tradition übersehen wird. Die Kopfnoten in der DDR zielten in einer Art von summativer Bewertung und in beträchtlicher Generalisierung auf „das für den Schüler typische Verhalten im Unterricht, in der produktiven Arbeit und in außerunterrichtlichen Veranstaltungen“. Erst der Beginn des Schuljahrs 1990/91 brachte in den neuen deutschen Bundesländern eine Änderung dieser Beurteilungspraxis.

3. Das Arbeits- und Sozialverhalten in der aktuellen Zeugnisreform

Bundesländer wie Brandenburg oder Thüringen, die derzeit dem eingangs skizzierten Drängen der Öffentlichkeit auf Erweiterung der Zensierung nachgeben, suchen einen stärker wissenschaftlichen Zugang und entwickeln mehr oder minder differenzierte Ratingskalen zur Einschätzung des „Arbeits- und Sozialverhaltens“ (s. Tabelle 1). Um überhöhten Erwartungen an die Präzision dieser Beurteilungen ebenso vorzubeugen wie deren pädagogisch einseitiger und Konkurrenz stiftender Interpretation, werden Skalenausprägungen genutzt, die sich von der Stufung der Ziffernnoten deutlich unterscheiden (stark, deutlich, ansatzweise, kaum). Damit nähern sich diese Beurteilungen jenen lernprozessorientierten Angaben, die – in lernförderlicher Absicht formuliert – zentrale Bestandteile von Lernentwicklungsberichten sind.

Inhaltlich betrachtet scheint derzeit eine erneute Phase der Konvergenz pädagogisch kontroverser Konzepte der schulischen Leistungsbeurteilung begonnen zu haben. Aus diagnostischer Sicht ist jedoch völlig ungeklärt, ob Lehrer dasjenige, was sie nun zusätzlich beurteilen und im Zeugnis zertifizierend feststellen sollen, auch hinreichend zuverlässig (d.h. objektiv und reliabel) und vor allem hinreichend zutreffend (d.h. valide) bewerkstelligen können. Empirische Untersuchungen zur Praxis der Kopfnotengebung gehen diesen Entscheidungen nicht voraus. Es bleibt zu hoffen, dass wenigstens prozessbegleitend eine solche Evaluation vorgenommen wird. Zudem muss aus pädagogisch-psychologischer Sicht angenommen werden, dass zumindest Lerneinstellungen (Interesse) sowie Lernstrategien und Selbsteinschätzungen der eigenen Kompetenz in erheblichem Maße bereichsspezifisch sind und deshalb als fachübergreifende Merkmale wenig praktische Evidenz besitzen.

Tabelle 1: Beispiel-Items aus der Zeugnisergänzung zum Arbeits- und Sozialverhalten (Entwurf von Brandenburg)

Informationen zum Arbeits- und Sozialverhalten				
	Ausprägung			
Merkmale	stark	deutlich	ansatzweise	kaum
(1) Selbstständigkeit: <i>löst Aufgaben selbstständig bedient sich notwendiger Hilfsmittel bzw. Medien kontrolliert die Ergebnisse selbst kann Informationen beschaffen, auswählen und gezielt verarbeiten ...</i>				
(2) Konfliktfähigkeit: <i>erkennt Interessengegensätze an respektiert Meinungen, Leistungen und Verhaltensweisen anderer trägt zur Konfliktlösung bei verhält sich im Auftreten angemessen ...</i>				

Wenn wie im Entwurf *Brandenburgs* (s. Tabelle 1) Einschätzungen zu (1) Prozessmerkmalen unterrichtlichen Lernens und zur (2) Kooperationsfähigkeit im Gruppenunterricht vorgenommen werden, dann zielt diese Beurteilung auf konkrete unterrichtliche Verhaltensweisen der Schüler und damit auf operationalisierbare Merkmale, die in der Unterrichtsplanung prinzipiell ausgewiesen werden können und denen konkrete Lerngelegenheiten im Unterricht entsprechen. Insofern werden hier „Ergebnisse“ schulischen Unterrichts erfasst, die – im engeren und klaren Sinne – als (1) „Arbeitsverhalten“ und (2) „Sozialverhalten“ bezeichnet werden können.

Sollen jedoch allgemeine Fähigkeiten von Schülern eingeschätzt werden, die auch im außerschulischen Bereich stark gefördert werden, so entstehen erhebliche diagnostische und pädagogische Probleme. Unstrittig gehören „Kommunikationsfähigkeit“ und „Konfliktfähigkeit“ oder auch „Toleranz“ zu den Grundfähigkeiten von Kindern und Jugendlichen. Aber schon die Beobachtung dieser Merkmale im Unterricht ist recht schwierig. Heikle Fragen können hier gestellt werden: Kann Konfliktfähigkeit anders als in Konflikten gezeigt werden? Müssen im Unterricht, wenn Konflikte ausbleiben, diese schon aus Beobachtungsgründen inszeniert werden? Sind jene Schüler konfliktfähiger, die unter diesen Bedingungen Konflikte inszenieren, oder jene, die deren Entstehung und deren Inszenierung vermeiden?

Die in Brandenburg bereits zum April 2000 erlassenen Verwaltungsvorschriften zu Informationen über das Arbeits- und Sozialverhalten (Ministerium für Bildung, Jugend und Sport 2000) weichen deutlich von der erprobten Form der differenzierten Schätzskalen ab. Nunmehr sind fünf Kategorien des Arbeits-

und Sozialverhaltens anhand vorgegebener, kategorienspezifischer Kriterien einzuschätzen: (1) Lerneinstellung, (2) Zuverlässigkeit, (3) Selbstständigkeit, (4) Teamfähigkeit, (5) Urteilsfähigkeit.

Am Beispiel der Kategorie „Selbstständigkeit“ soll das gewählte Verfahren verdeutlicht werden. Folgende Kriterien werden hier vorgegeben:

- a) arbeitet sehr selbstständig und sehr flexibel sowie mit hoher Eigeninitiative,
- b) arbeitet selbstständig und findet oft eigene Lösungsansätze, entwickelt eigene Ideen,
- c) arbeitet mit kleinen Anregungen weitgehend selbstständig, zeigt Eigenaktivität,
- d) bei beständiger Hilfe kann die Aufgabenstellung selbstständig erfüllt werden,
- e) arbeitet noch zu selten selbstständig und benötigt zusätzliche Anleitung.

Schon die verbale Beschreibung der Kategorien lässt vermuten, dass unterschiedliche Aspekte des Arbeitsverhaltens in wechselnder Kombination beurteilt werden sollen. So wird das kategorienbestimmende Attribut „selbstständig“ über die Stufen hinweg in uneinheitlicher Weise präzisiert. Auf Stufe a) soll unter anderem dessen *Intensität* („sehr“) beurteilt werden, während auf Stufe e) dessen *Frequenz* („zu selten“) einzuschätzen ist. Die Vermischung von Intensitäts- und Frequenzmerkmalen in einer Skala stellt ein ernstes diagnostisches Problem dar, das sich auch in den Skalierungsbemühungen sog. Rasterzeugnisse (vgl. Garlichs 2000; Arnold 1997) zeigt. Ob die Spezifizierungen „mit kleinen Anregungen“ (c), „bei beständiger Hilfe“ (d) und „benötigt zusätzliche Anleitung“ eine hinreichende klare und damit konsensuell handhabbare Stufung des Merkmals Selbstständigkeit ermöglichen, scheint zumindest für die beiden letztgenannten sehr fragwürdig. Wird „beständige Hilfe“ nicht dann gegeben, wenn „zusätzliche Anleitung“ bereits nicht mehr hinreichend ist? Die zusätzlich gegebene definitorische Setzung, dass das übergreifende Merkmal Selbstständigkeit aus den Einzelmerkmalen „Eigeninitiative, Arbeitsweise, Arbeitstempo, Selbstkontrolle, Kommunikationsfähigkeit, Einschätzen der eigenen Leistungsfähigkeit“ bestehen soll, lässt vermuten, dass den Lehrern die Beurteilung eines zumindest sehr komplexen Konstruktes vorgeschrieben wird, ohne dass eine hinreichend präzise und obendrein wenig stimmige Operationalisierung verfügbar ist.

Und selbst wenn Lehrer aufgrund ihrer in der Berufspraxis erworbenen Kompetenz diese beträchtliche Beurteilungsleistung hinreichend konsistent erbringen, so bleibt dennoch die Frage offen, worüber sich Schüler und Eltern mit diesen Beurteilungen informiert fühlen können. Wenn die Eltern im Zeugnisformular zu jeder Beurteilungskategorie nur das verbal ausformulierte Kriterium vorfinden (für „Ergänzende Informationen“ ist reichlich Textraum vorgesehen), werden ihnen die zur Interpretation notwendigen Bezugsinformationen, d.h. die übrigen vier Kriterienstufen, nicht angeboten. So muss für sie unklar bleiben, ob ihr Kind mit einer anderen Kriterienstufe zutreffender beurteilt würde oder mit keiner der vorgegebenen zutreffend beurteilt werden kann.

Bei aller Kritik muss gleichwohl darauf hingewiesen werden, dass die Brandenburgische Verordnung trotz der problematischen Beurteilungskriterien

deutlich differenzierter angelegt ist als Beurteilungserlasse einiger anderer Bundesländer. So ist im *Saarland* zum 19. April 2000 die „Verordnung über Verhaltenszeugnisse“ (Amtsblatt S. 828) in Kraft getreten, die eine Ergänzung der Abschluss- bzw. Abgangszeugnisse um eine vierstufige Zensierung der Merkmale „Betragen, Mitarbeit, Arbeitshaltung und Teamfähigkeit“ festlegt. Der Runderlass des *Niedersächsischen* Kultusministeriums vom 8.3.2000 sieht eine vierstufige Zeugnisbewertung von „Arbeitsverhalten“ und „Sozialverhalten“ vor und bietet den Gesamtkonferenzen der Schulen die Möglichkeit, sich für eine Ergänzung oder den Ersatz der „standardisierten Form“ durch verbale Beurteilungen zu entscheiden.

Bereits im Schuljahr 1999/2000 führte der Freistaat *Sachsen* die Kopfnoten auf den Zeugnissen wieder ein und knüpfte damit an die Beurteilungspraxis der früheren DDR ebenso an wie an jene zwei Jahrzehnte zuvor in der Bundesrepublik aufgegebene Zeugnisform. In seiner Begründung bezog sich der sächsische Kultusminister auf das mehrheitliche Votum der Eltern für diese Form der Zensierung (vgl. Rössler 2000). „Ordnung, Mitarbeit, Betragen und Fleiß“ sind mit den Noten 1 bis 5 zu bewerten, was zunächst die hinlänglich bekannte Frage nach gültigen Bewertungskriterien provoziert: Auf der Grundlage welcher Fakten und mit welchen Methoden können Lehrer einschätzen, ob Betragen, Fleiß, Mitarbeit oder Ordnung eines Kindes *vorbildlich* (Note 1), *stark* (Note 2), *durchschnittlich* (Note 3), *schwach* (Note 4) oder *unzureichend* (Note 5) *ausgeprägt* sind? Das Problem spitzt sich dramatisch zu, wenn man sich anschaut, was sich nach Vorstellungen des Kultusministeriums hinter diesen recht tradierten Begriffen verbergen soll. In einem Elternbrief heißt es z.B., dass *Mitarbeit* die Initiative, Kooperationsbereitschaft und Teamfähigkeit, Beteiligung im Unterricht, Selbstständigkeit, Kreativität und Verantwortungsbereitschaft des Kindes umfassen soll. Die Note für das *Betragen* soll dann die Aufmerksamkeit, Hilfsbereitschaft, Zivilcourage und angemessenen Umgang mit Konflikten, Rücksichtnahme, Toleranz, Gemeinsinn und Selbsteinschätzung eines Kindes bewerten. In ähnlicher Weise soll auch die Bewertung von Fleiß und Ordnung diverse Persönlichkeitseigenschaften in einer Note zusammenfassen (vgl. Sächsisches Staatsministerium für Kultus 1999, S. 2).

Für diese Zusammenfassungen sehr unterschiedlicher Verhaltens- und vor allem Persönlichkeitsmerkmale gibt es bislang weder theoretische Modelle noch empirisch gesicherte Verfahren. Insofern muss angenommen werden, dass die beurteilenden Lehrer kaum über die erforderlichen diagnostische Kompetenzen verfügen. Generalisierungen und Kategorisierungen oben genannter Art sind in der Sozialpsychologie Gegenstand der Stereotyp- und der Vorurteilsforschung. Hofer (1969) hat dafür den Begriff der „Impliziten Persönlichkeitstheorien von Lehrern“ vorgeschlagen, zu deren Aufklärung und Differenzierung pädagogische Diagnostik beitragen sollte, um zugleich deren Herausbildung und Habitualisierung entgegenzuwirken.

4. Sozialerzieherische Wirkung und moralische Rechtfertigung

Die so genannten Kopfzensuren haben mit den Merkmalen des Arbeits- und Sozialverhaltens eine gewisse Schnittmenge: zu einem beträchtlichen Teil bilden sie eine unentwirrbare Mischung aus verhaltensbezogenen Merkmalen der Schüler und schul- sowie unterrichts- und damit lehrerbezogenen Einstellun-

gen. Ob es in einer demokratischen Schule jedoch legitim und pädagogisch erwünscht ist, die Ausbildung von Einstellungen zum Gegenstand von Beurteilungen zu machen, das ist keineswegs klar. Viele Argumente sprechen dagegen, „Gesinnungsprüfungen“ dieser Art vorzunehmen.

Und selbst wenn diese Argumente mit pädagogischer Aufrichtigkeit entkräftet würden, so bleibt dennoch das entscheidende methodische Grundproblem ungelöst: die „Regelorientierung“ der Schüler (Pünktlichkeit, Zuverlässigkeit, Beteiligung am Unterricht, Zivilcourage, Rücksichtnahme, Zielstrebigkeit, Ausdauer etc.) hängt in großem Maße davon ab, ob diese Merkmale auch auf den erteilten Unterricht zutreffen. Was bedeutet Selbstständigkeit in einem eng geführten Frontalunterricht? Sind Schüler tolerant, die ihre Lehrerin oder ihren Lehrer auch dann nicht kritisieren, wenn sie oder er Mitschüler als „Schwachmatiker“ oder „Chaoten“ titulieren? Sind hausaufgabenerledigende Schüler nicht insbesondere dann aner kennenswert fleißig, wenn Unterrichtsdefizite auszugleichen sind, zu denen auch ihre Lehrer oder ihre Lehrerinnen beigetragen haben?

Das Thema „Kopfnote“ bzw. „Beurteilung des Arbeits- und Sozialverhaltens“ tangiert letztlich die Frage, in welcher Weise es pädagogisch sinnvoll und diagnostisch möglich sein kann, Erziehungswirkungen des schulischen Unterrichts zu erfassen. Dass Schule nicht nur fachlich qualifizieren, sondern auch bilden und zur Gemeinschaftlichkeit erziehen soll, spiegelt sich weder in den traditionellen Zeugnissen noch in den systembezogenen Evaluationsstudien wider. Beide Fakten gründen in derselben Problematik: Unterricht und damit Schule erzieht zeitgleich mit der familiären Situation und dem Milieu der Freizeitaktivitäten, wobei nur zu einem Teil gleiche Erziehungsziele verfolgt und beträchtliche Diskrepanzen entweder akzeptiert, toleriert, nicht verhindert oder konterkariert werden. Jedenfalls ist eine Bewertung der schulischen Erziehungswirkungen nur unter der Voraussetzung möglich, dass alle drei genannten Wirkungsbereiche erfasst werden – und selbst dann bleiben noch beträchtliche methodische Probleme ungelöst. Man stelle sich deshalb vor, dass zur Zeugniserstellung auch ein Elternfragebogen über die häusliche Erziehungssituation und über das Freizeitverhalten der Schüler ausgeteilt und von den Lehrern ausgewertet würde. Kann und darf Schule diese Auskünfte – regelmäßig – verlangen? Dürfen Zeugnisse „Nettowirkungen“ ihrer erzieherischen Bemühungen berichten, d.h. Erfolge, die sich „trotz dem“, d.h. den problematischen familiären Verhältnissen trotzend, eingestellt haben? Oder, um die selten beachtete Kehrseite zu zeigen: Kann Schule auch dann sehr gute Erziehungswirkungen in den Kopfnoten bescheinigen, wenn Kinder aus bildungsinteressierten Elternhäusern daheim zu vielseitigen Lernaktivitäten angeregt werden und ihr Sozialverhalten in vollständigen Familien ohne Knappheit an jedweden Konsumgütern von akademisch verbalisierenden Eltern einfühlend geprägt ist? Vielleicht erhalten sie aber auch nur eine „2“ in „Mitarbeit“, weil sie etwas schüchtern oder zurückhaltend sind, wenn alle anderen sich schon lärmend melden?

5. Kontextgebundenheit der Beurteilung

Das Sozialverhalten sowie die unterrichts-, lehrer- und schulbezogenen Einstellungen können kaum anders als „kontextbezogen“ beurteilt werden. In der

Terminologie der pädagogisch-psychologischen Diagnostik bedeutet dies: diese Beurteilungen sind „sozialnormorientiert“ und gelten – zunächst, d.h. ohne weitere empirische Evidenz – nur für eine konkrete Konstellation, d.h. für diese Schulklasse, deren Lehrer und die von den Lehrern realisierte Unterrichtsqualität. Sachnormorientierte Beurteilungen hingegen sind hier nur unter Rückgriff auf sehr stark generalisierende und damit pädagogisch wenig aussagekräftige Konstrukte möglich.

Dieser Einwand richtet sich gegen eine Beurteilung des Sozialverhaltens und der unterrichtsbezogenen Einstellungen, wenn die „Bezugsnorm“ nicht mitgeteilt wird, d.h. wenn die Unterrichtsbedingungen als Information und als Bezugsrahmen unerwähnt bleiben. Auch aus diagnostisch besten Gründen umfasst das Zeugnis-konzept der Laborschule Bielefeld deshalb den „Unterrichtsbericht“. In dieser Weise ist das Kontextualisierungsproblem der Beurteilung des Arbeits- und Sozialverhaltens durchaus lösbar: der Bezugsrahmen der sozialen Merkmale der Schulklasse und der Merkmale der Unterrichtsqualität wird mitgeteilt und nachfolgend die Bindung des Arbeits- und Sozialverhaltens an diese situativen Bedingungen.

6. Hinweise auf die Unsicherheit der Verhaltensbeurteilung

Zu den Grundeigenschaften wissenschaftlich begründeter pädagogischer Diagnostik gehört sowohl die hinreichende Präzisierung der Beurteilungen als auch die deutliche Anzeige der Ungenauigkeit, die unvermeidlich mit ihnen verbunden ist. Präzisierung bedeutet, dass der spezifische unterrichtliche Kontext erläutert wird, in dem Erziehungsziele erreicht oder verfehlt werden. Dies ist kaum anders möglich als durch beschreibende Angaben, die jedenfalls klassenübergreifend kaum zu standardisieren sind. Die Mitteilung von Unsicherheit in den Beurteilungen führt in ihrer „verdeckten“ Form dazu, dass Lehrer „Betragen“ und „Fleiß“ offenbar deutlich milder beurteilen als fachliche Leistungen (vgl. Weiß 1989, S. 101).

Offen wird mit dieser Unsicherheit umgegangen, wenn die Zeugnistexte diese Vagheit verbal oder durch Intervalle, in denen die zutreffende Einschätzung vermutlich liegt, anzeigen. Diagnostisch gesehen stellt die Mitteilung der inhärenten Unsicherheit von Beurteilungen nicht nur einen methodischen Standard dar. Hier zeigt sich auch der Praxisbezug diagnostischer Entscheidungen: diese sind nur dann effizient, wenn Klarheit und somit keine Spekulation über die Ungenauigkeit der Beurteilungsprozesse besteht. So kann es durchaus Konstellationen geben, in denen relativ ungenaue Beurteilungen dennoch hohe Entscheidungseffizienz bedingen, z.B. wenn das in Frage stehende Verhalten bei Schülern fast nie oder sehr selten auftritt. Es kommt also für die Beurteilung des Sozialverhaltens von Schüler X nicht so sehr darauf an, ob er zwei-, drei- oder viermal seinen afrikanischen Mitschüler aus der Horde der ihn prügelnden Mitschüler befreit hat.

Jedoch: Zumeist richtet sich die Beurteilung des Arbeits- und Sozialverhaltens auf relativ häufig auftretende Merkmale des Schülerverhaltens. Und just hier besteht sowohl eine erhebliche Beurteilungsunsicherheit als auch die Gefahr der subjektiv verzerrenden Perspektive, was in positiver Formulierung die „Subjektivität jeglicher Leistung und Vergleichbarkeit“ meint (Tillmann/Vollstädt 1999, S. 45). Diese menschlich, allzu menschliche Vagheit des Urteilens vor

den Schülern und deren Eltern zu kaschieren, bedeutet auch, sich der rationalen Begründung eigenen Beurteilens zu enthalten, was an jene Geisteshaltung erinnert, gegen die in Deutschland und Europa eine kulturhistorisch einmalige Bewegung vor drei Jahrhunderten angetreten ist, deren Früchte wir meinen für unsere moderne Gesellschaft ernten zu dürfen: die Aufklärung. Der hier zum tragenden Erziehungs- und Staatsprinzip erhobene Appell an den kritischen Gebrauch der eigenen Vernunft gilt auch den sozialen Verhältnissen und Einrichtungen der Menschen. Die Beurteilungspraktiken in unseren Schulen sollten vor der aufklärenden Vernunft bestehen können und somit ihre Grenzen nicht verschweigen.

Literatur

- Arnold, Karl-Heinz 1997: Formen und Inhalte von Rasterzeugnissen: Eine Untersuchung der strukturierten Berichtszeugnisformulare in den Grundschulen des Bundeslandes Bremen (verfügbar unter: <http://www.tu-berlin.de/fb2/sowi/paedpsy/mat/RaZeug0101.pdf>).
- Arnold, Karl-Heinz / Jürgens, Eiko 2001: Schülerbeurteilung ohne Zensuren. Neuwied: Luchterhand
- Döbert, Hans / Geißler, Gert 2000: Schulleistung in der DDR: das System der Leistungsentwicklung, Leistungssicherung und Leistungsmessung. Frankfurt am Main: Lang
- Dohse, Walter 1967: Das Schulzeugnis: sein Wesen und seine Problematik. 2. Aufl. Weinheim: Beltz
- Garlichs, Ariane 2000: Kennen Sie Rasterzeugnisse? Eine Empfehlung zur Nachahmung. In: Grundschulzeitschrift. 14, 2000, 135/136, S. 78-79 (Anhang dazu S. 52-58)
- Hofer, Manfred 1969: Die Schülerpersönlichkeit im Urteil des Lehrers. Eine dimensionsanalytische Untersuchung zur impliziten Persönlichkeitstheorie. Weinheim: Beltz
- Ingenkamp, Karlheinz (Hg.) 1974: Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte. 7. Aufl. Weinheim: Beltz
- Ingenkamp, Karlheinz / Laux, Hermann 1990: Geschichte der pädagogischen Diagnostik. 2 Bde. Weinheim: Deutscher Studienverlag
- Lehmann, Jens / Ziegenspeck, Jörg W. 2000: Leistung und/oder Disziplin? Konstruktiv-kritische Anmerkungen zur Diskussion über die Wiedereinführung von „Kopfnoten“. Die Deutsche Schule. 92, 2000, 2, S. 218-223
- Ministerium für Bildung, Jugend und Sport des Landes Brandenburg 2000. Verwaltungsvorschriften zu Informationen über das Arbeits- und Sozialverhalten (VVarb-Soz) vom 17. April 2000. Amtsblatt des MfBJS, 9. Jahrgang, Nummer 6, S. 202-207
- Ministerium für Volksbildung der DDR 1980: Zeugniserteilung, Versetzungsbestimmungen, Gesamteinschätzung und Verhaltenszensurierung. Anweisung über die Erteilung von Zeugnissen und die Versetzung an den allgemeinbildenden polytechnischen Oberschulen vom 8. Januar 1980. Berlin (Ost): Volk und Wissen
- Rauschenberger, Hans 1999: Umgang mit Schulzensuren. Funktionen – Entwicklung – Praxis. In: Grünig, B. u.a. (Hg.): Leistung und Kontrolle. Weinheim: Juventa, S. 11-100
- Rössler, Matthias 2000: Kopfnoten als Erziehungsmittel? In: Pädagogik. 52, 2000, 10, S. 60
- Sächsisches Staatsministerium für Kultus 1999. Elternbrief 11. Dresden
- Tillmann, Klaus-Jürgen / Vollstädt, Witlof 1999: Funktionen der Leistungsbewertung: Eine Bestandsaufnahme. In: Pädagogik. 51, 1999, 2, S. 42-46
- Weiss, Rudolf 1989: Leistungsbeurteilung in den Schulen – Notwendigkeit oder Übel? Problemanalysen und Verbesserungsvorschläge. Wien: Jugend und Volk

Witzlack, Gerhard 1986: Verhaltensbewertung und Schülerbeurteilung. 3. Aufl. Berlin (Ost): Volk und Wissen

Karl-Heinz Arnold, geb. 1952, Dr. phil. et habil., Dipl.Psych., Dipl.Päd., Erstes Staatsexamen für das Lehramt mit Wahlfach Mathematik, Professor für Pädagogische Psychologie an der Technischen Universität Berlin

Anschrift (privat): Kielmoorweg 23, 27574 Bremerhaven

E-Mail: karnold@tu-berlin.de

Witlof Vollstädt, geb. 1945, Prof. Dr. paed. habil; wiss. Mitarbeiter an der Gesamthochschule Kassel

Anschrift (privat): Gerhart-Hauptmann-Str. 6, 34281 Gudensberg

E-Mail: Witlof.Vollstaedt@t-online.de