

Kalthoff, Herbert 1996: Das Zensurenpanoptikum. Eine ethnographische Studie zur schulischen Bewertungspraxis.
Zeitschrift für Soziologie, Jahrgang 25, Heft 2, S. 106 - 124

Die Zeitschrift wird herausgegeben von der UNIVERSITÄT BIELEFELD, FAKULTÄT FÜR SOZIOLOGIE. Die Fakultät beruht auf Zeit, das in seiner Tätigkeit unabhängige Herausbergremium und den Beirat.

Herausgeber:

JOHANNES BERGER, Mannheim
JOHANN HANDL, Mannheim
STEFAN HIRSCHAUER, Bielefeld
CHRISTEL HOPE, Hildesheim
WOLFGANG STREECK, Köln

Redaktion und

STEFAN HIRSCHAUER
Universität Bielefeld, Fakultät für Soziologie, Postfach 10 01 31, D-33501 Bielefeld
Telefon (05 21) 106-46 27 (vorm. Sekr. -46 26) · Fax (05 21) 106-58 44
e-mail: stefan.hirschauer@post.uni-bielefeld.de

Wissenschaftlicher Beirat:

JUTTA ALLMENDINGER, München · BERNHARD BADURA, Bielefeld · JORG BERGMANN, Gießen · LARS CLAUSEN, Kiel
ANDREAS DIEKMANN, Birm · HANS DIETER EVERS, Bielefeld · UTE GERHARD, Frankfurt · MICHAEL T. GREVEN, Darmstadt ·
HARTMUT HAUSERMANN, Berlin · AXEL HONNETT, Berlin · HANS JOAS, Berlin · FRANZ-XAVER KAUFMANN, Bielefeld ·
HANSRIED KELLNER, Frankfurt · HANS-DIETER KLINGEMANN, Berlin · KARIN KNORR CETINA, Bielefeld · MARTIN KOHL, Ber-
lin · ROlf LINDNER, Berlin · THOMAS LUCKMANN, Konstanz · KURT LÜSCHER, Konstanz · NIKLAS LUHMANN, Bielefeld ·
BURKART LUTZ, München · BIRGIT MAHNKOPF, Berlin · KARL ULRICH MAYER, Berlin · RENATE MAVNTZ, Köln · WALTER
MÜLLER, Mannheim · RICHARD MÜNCH, Bamberg · ROSEMARIE NAVE-HEINZ, Oldenburg · GERTRUD NUNNER-WINKLER, Mün-
chen · ILONA OSTNER, Göttingen · FRANZ U. PAPP, Mannheim · DETLEF POLLACK, Leipzig · WERNER RAMMERT, Berlin ·
OTTHEIN RAMNSTEDT, Bielefeld · MARIANNE RODENSTEIN, Frankfurt · GERT SCHMIDT, Erlangen · YVONNE SCHÜTZE, Berlin ·
HANS-GEORG SOEFFNER, Konstanz · ILVA SRUBAR, Erlangen · GUNTHER TEUBNER, London · PETER WEINGART, Bielefeld ·
WOLFGANG ZAFF, Berlin

Die Zeitschrift veröffentlicht Beiträge aus allen Bereichen der Soziologie und ihren Randgebieten einschließlich methodologischer und forschungstechnischer Arbeiten. Beiträge, die sowohl einen theoretischen als auch einen empirischen Bezug aufweisen, werden besonders begrüßt. Kritische Kommentare (maximal 5 Schreibmaschinenseiten) zu einzelnen in der Zeitschrift erschienenen Beiträgen sind erwünscht und können zusammen mit Erwidrerungen der betroffenen Autoren abgedruckt werden. Die Annahme zugegangener Manuskripte bleibt vorbehalten. – Es ist vorgesehen, fallweise Literatur- und Besprechungsaufsätze, nicht jedoch Einzelrezensionen aufzunehmen. Rücksendungen unverlangter Bücher können nicht vorgenommen werden.

Mit der Annahme eines Manuskripts erwirbt der Verlag für die Dauer der gesetzlichen Schutzfrist die ausschließliche Befugnis zur Wahrnehmung der Verwertungsrechte im Sinne der §§ 15 ff. des Urheberrechtsgesetzes. Die Zeitschrift und alle in ihr enthaltenen einzelnen Beiträge und Abbildungen sind für die Dauer des Urheberrechts geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. – Die Aufnahme der Zeitschrift in Lesezirkel ist nicht gestattet.

© 1996 Ferdinand Enke Verlag, Bludenzter Straße 6, D-70469 Stuttgart · Printed in Germany · Satz: Photocomposition Jung, F-67420 Diesbach/Plaine, Druck: Druckhaus Götz GmbH, D-71636 Ludwigsburg, Verantwortlich für den Textteil: Stefan Hirschauer, Universität Bielefeld, Fakultät für Soziologie, Postfach 10 01 31, D-33501 Bielefeld; für den Anzeigenteil: Ferdinand Enke Verlag, Postfach 30 03 66, D-70443 Stuttgart, Telefon (07 11) 13 57 98 20, Fax 13 57 98 30. Erscheinungsweise: Jährlich 1 Band zu 6 Hefen (zweimonatlich). Bezugspreis jährlich DM 142,- (unverbindlich empfohlener Preis); für Studenten gegen Vorlage einer Bescheinigung DM 69,- zuzüglich Versandkosten (Inland: DM 12,60, Ausland: DM 16,80); Einzelheft DM 30,- zuzüglich Versandkosten. Erlöge bis zum 30. September keine Abbestellung. verhängt sich die Bezugsdauer jeweils um ein Jahr. Rechtsinhaber: Neubesitzungen können innerhalb von 8 Tagen beim Verlag, Ferdinand Enke Verlag, Postfach 30 03 66, D-70443 Stuttgart, Telefon (07 11) 89 31-0, Fax 8 93 12 98) widerrufen werden. Es genügt die rechtzeitige Absendung der Widerrufserklärung (Poststempel).

Die Zeitschrift für Soziologie wird regelmäßig von „Bulletin Signalétique de Sociologie“, „Current Contents“, „Social Science Citation Index“, „Sociological Abstracts“, „Sociology of Education Abstracts“, „Psychologischer Index“, „Sozialwissenschaftliches Literaturinformationssystem SOLIS“ und „Zeitschriftendaten“ des Deutschen Bibliotheksinstitutes erfüllt.

For users in the USA

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Ferdinand Enke Verlag Stuttgart for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the base fee of \$ 7.00 per copy of each article is paid directly to CCC, 21 Congress St., Salem, MA, 01970. 0340-1804/96 \$ 7.00.

Inhalt

- | | | |
|-----|---|------------------|
| 89 | Theorie
Differenzierung als absoluter Begriff?
Zur Revision einer soziologischen Kategorie | Gerhard Wagner |
| 106 | Kultursozologie
Das Zensurenparadoxium. Eine ethnographische Studie zur schulischen Bewertungspraxis | Herbert Kalthoff |
| 125 | Wirtschaftssoziologie
Was ist soziologisch an der Wirtschaftssoziologie?
Ungewißheit und die Einbettung wirtschaftlichen Handelns | Jens Beckert |
| 147 | Entscheidungstheorie
Akzeptanzbeschaffung: Verfahren und Verhandlungen | Hendrik Vollmer |
| 165 | Mitteilungen
„Work Quo Vadis?“ International Conference at the University of Karlstad, Sweden, 15.-17.6.96 | |
| 165 | Preis der Fritz-Thyssen-Stiftung für Zeitschriftenaufsätze 1994 | |
| 168 | Die Autoren dieses Heftes | |
| 169 | English Contents Page | |
| 170 | English Abstracts | |

Das Zensurenpanoptikum Eine ethnographische Studie zur schulischen Bewertungspraxis¹

Herbert Kalthoff

Europa-Universität Viadrina, Frankfurter Institut für Transformationsstudien, Postfach 776, D-15207 Frankfurt/Oder

Zusammenfassung: Die institutionelle Fiktion der Schule ist es, daß die Leistung von Schülern ihre Bewertung konstituiert. Der Beitrag geht hinter diese Annahme zurück und gibt einen mikroanalytischen Einblick in die Bewertungspraxis von Lehrpersonen. Er konzeptualisiert seine theoretische Perspektive in Auseinandersetzung mit der Bewertungsforschung und einer empirischen Studie Bourdieus. Gegen die Normativität der Bewertungsforschung und gegen den reproduktiven Bias der Bourdieuschen Praxistheorie entläßt der Beitrag auf der Basis von ethnographisch gewonnenen Daten und dem empirischen Programm des Konstruktivismus eine rekontextualisierende Perspektive. Er zeigt, daß im Bewertungsvollzug die Bewertung einer Schulleistung und Selbstbeobachtung der Lehrperson miteinander gekoppelt sind. Das Phänomen der Involvierung verweist darauf, daß die Bewertung mehr ist als eine Abbildung von Meßergebnissen, sondern daß Lehrpersonen immer auch ihre eigene Leistung konstruieren. Dies wird an zwei Formen der Beobachtung dargestellt: Selbstbeobachtung im Kontext der Korrektur schriftlicher Klausurarbeiten und Fremdbeobachtung im Kontext der mündlichen Abiturprüfung.

1. Einleitung

Es ist für uns, die wir die Schule durchlauten haben, selbsterfindlich, daß Lehrpersonen das Wissen ihrer Schüler schriftlich oder mündlich überprüfen, und daß sie das, was sie die „Leistung“ der Schüler nennen, mit einer Note bewerten. Die Institution Schule ist nicht der einzige gesellschaftliche Ort, an dem Leistung bewertet wird. Man denke etwa an die Punkt- und Preisrichter im Sport; auch im Wissenschaftssystem wird bewertet, etwa bei der Auswahl von Manuskripten einer Zeitschrift.

Soziologisch interessant ist die schulische Bewertungspraxis in zweifacher Hinsicht: Zum einen vermittelt die Schule Wissen an ihre Schüler, zum anderen organisiert sie in Form von Zensuren und Abschlüssen gesellschaftlich relevantes Wissen über Schüler und macht sie miteinander vergleichbar. Schüler erhalten am Ende ihrer Schulbahn ein Zertifikat: eine *Acta Paß*, der über die Schulzeit – über gewählte Kurse, über Stärken und Schwächen – Auskunft erteilt und die Zutrittschancen zur professionellen Laufbahn regelt. Und es gehört zum Standard der Sozialstrukturanalyse, diese Bildungszertifikate als unabhängige Variable der Statuszuweisung zu betrachten (vgl. Geißler 1987).

¹ Stefan Hirschnauer danke ich ausdrücklich für die intensive Diskussion bei der Abfassung des Artikels, Klaus Amann 'for all his personal comments', den Herausgebern sowie der Gutachterin für Anregungen und Kritik.

Ziel dieses Aufsatzes ist es, einen Beitrag zu einer kultursociologischen Analyse der schulischen Bewertungspraxis zu leisten, d. h. einen Blick in den Erzeugungsprozeß dessen zu werfen, was die Sozialstrukturanalyse als ein 'Faktum' betrachtet. Im Rahmen der Kultursociologie existieren nur wenige Arbeiten, die empirisch die Bewertungspraxis von Lehrpersonen untersucht haben. Verschiedene Untersuchungen betonen hingegen den Zusammenhang von Bewertungsergebnissen (d. h. schulischem Erfolg) und Konservierung der Sozialstruktur entwickelter Gesellschaften. Diese reproduktionstheoretisch angelegten Studien unterstreichen die Bedeutung des Schulsystems für die Reproduktion klassenspezifischer Differenzen (vgl. z. B. Bourdieu/Passeron 1971; Apple 1982; Collins 1979; Barton/Walker 1983; Cookson/Persell 1985). Das Reproduktionsmodell betont insbesondere die Relation von Ökonomie und Schulsystem. Die Schule vermittelt danach die spezifischen Kenntnisse und Fähigkeiten, die für den späteren Arbeitsplatz erforderlich sind; auf diese Weise stabilisiert sie die gesellschaftliche Arbeitsteilung (vgl. Giroux 1983). Auf die Bewertungspraxis von Lehrpersonen aber, die das Gelesene auf die eine oder andere Art attestieren, gehen diese Studien nicht ein.

Die Theoretiker kultureller Reproduktion haben insbesondere die Bedeutung der Schule für die Legitimierung der dominanten Kultur in allen ihren Äußerungsformen betont; hier sind vor allem die Arbeiten Bourdieus zu nennen (Bourdieu 1989, 1984).

In einer kleinen empirischen Studie analysieren Bourdieu/Saint Martin (Bourdieu/Saint Martin 1975; Bourdieu 1989: 48 ff.) 154 Karteikarten, auf die der Lehrer einer Vorbereitungsklasse für die Aufnahmeprüfung an eine Pariser Eliteschule neben persönlichen Daten der Schülernamen (bis heriger Schulbesuch, Beruf der Eltern etc.) fünf bis sechs Noten für schriftliche Arbeiten und mündliche Beteiligung sowie Beurteilungsgründe notiert hat. Von diesem Material ausgehend, entwerfen Bourdieu/Saint Martin das Bild einer komplexen „Bewertungsmaschine“, durch die hindurch die soziale Rangordnung der Schüler in eine schulische transformiert wird, d. h. das schulische Urteil objektiviert und verstärkt das (in-korporierte) kulturelle Wissen der sozial herrschenden.² Dabei spielen, so Bourdieu/Saint Martin, sachliche Gründe im Beurteilungsvorgang nur eine Nebenrolle; sachfremde Gründe hingegen die Bewertung. Die 'Maschine' vollbringt nicht nur eine Bestätigung der gesellschaftlichen Ordnung, sondern sorgt zugleich dafür, daß ihre Funktionsweise und ihre Funktion unferkannt bleiben. Als „ideologische Maschine“ mystifiziert sie sowohl die Ergebnisse als auch die Akteure: Anerkennung und Verkennung der schulischen Urteilskraft und ihrer sozialen Wirkungen – Attributions- und Reproduktionsfunktion – gehen zusammen. Die Reproduktion sozialer Positionen wird von Bourdieu/Saint Martin (1975: 76 f.) handlungstheoretisch durch einen objektiven Handlungszwang, durch den Glauben an den Sinn und den Zweck des eigenen Tuns sowie durch die Homologie von Hierarchie des Bildungssystems und mentalen Strukturen der Akteure, die wiederum selbst Produkt des Bildungssystems sind, erklärt.

Die Analyse der Bewertungspraxis ist bei Bourdieu eingebettet in verschiedene Theoriekontexte: Im Kontext der Theorie der pädagogischen Aktion, der Habitusstheorie sowie der Differenzierungstheorie der Schule wird die gesellschaftliche Funktion der Schule in der Perpetuierung sozialer Macht gesehen. Dem entsprechen bei Bourdieu drei zentrale Dimensionen: Durchsetzung kultureller Willkür mittels symbolischer Gewalt, unglei-

che Verteilung schulischer Titel (und somit ungleiche Erfolgchancen) und (körperliche) Systematisierung familiär erworbener Dispositionen, die zu einer Homogenisierung sozialer Klassen führt (vgl. Bourdieu 1989, 1979; 1974; Bourdieu/Passeron 1973; 1971).

Umfangreiche Forschungen zur schulischen Urteilspraxis finden sich dagegen in der Bewertungsforschung: in ihr überwiegt die Kritik an der herkömmlichen Schülerbeurteilung. Allerdings verweist das Alter der Untersuchungen darauf – einige Studien wurden Mitte der 1970er Jahre neu publiziert –, daß die Beurteilung von Schülern in der bildungssoziologischen Forschung, von Untersuchungen im angelsächsischen Raum abgesehen (vgl. Brookhart 1993; Stiggins/Conklin 1992), kaum mehr ein Thema ist. Mitte der 1980er Jahre löst eine Reevaluation des Lehrerturms (vgl. Sommer 1983; Steinhilf 1983; Schröder 1990) eine nach Jahrzehnten der Wiederholung verpuffte Kritik ab. Die „alte Klage“, der Bewertungsforschung bezieht sich dabei insbesondere auf die Verletzung des Gleichheitspostulats (vgl. Luhmann/Schorr 1988: 314 ff.). Eine Vielzahl von Untersuchungen hat zu zeigen versucht, daß die Ergebnisse des schulischen Bewertungsverfahrens nicht den Gütekriterien der Objektivität, Reliabilität und Validität standhalten. Versuchs kombinationen belegen empirisch die Ungenauigkeit des schulischen Bewertungsmittels „Zensur“ durch die Fehlerquelle und den Unsicherheitsfaktor „Lehrer“. Diese Studien zeigen zum einen, daß die in Bewertungsvorfahren involvierten Lehrpersonen unabhängig von der schulischen Disziplin weder bei der Beurteilung ein und derselben Klausurarbeit zum gleichen Ergebnis kommen,³ noch bei einer wiederholten Beurteilung den gleichen Klausurtext mit der zuvor gegebenen Note zensurieren (vgl. Eells 1977 [1930]; Finlayson 1977 [1951]; Starck/Elliot 1977 [1913]; Schröder 1971; Köter/Grau 1965; Härtog/Rhodes 1934). Zum anderen belegen die Untersuchungen, daß das Urteil der Bewertungspraktiker durch Informationen zur Person der Schüler – z. B. zur sozialen Herkunft, zum schulischen Leistungsstand oder zur zum Geschlecht – beeinflusst wird

² Ein Beispiel: Aus „Sie sind ein Arbeitersohn“ wird „Sie sind gewöhnlich“. Der schulischen Beurteilung wohnt die symbolische Kraft inne, beim Beurteilen als Zensur zu wirken: „Philosophie legt mir nicht“ oder „Ich liebe Geographie“ (Bourdieu/Saint Martin 1975: 80 f.; vgl. auch Bourdieu 1989: 163 ff.).

³ Zum Beispiel zeigt Weiss (1965: 148 ff.), daß 153 Lehrer dieselbe Mathematikarbeit mit Noten von „sehr gut“ (7%), „gut“ (4%), „befriedigend“ (42%), „ungünstig“ (9%) bis „nicht genügend“ (1%) bewerten. Unschöfer (1949) belegt Beurteilungsdifferenzen für das Fach Deutsch anhand von Abituraufätzen.

vgl. Wierczkowski/Nickel 1974; Ferdinand 1971; Knoche 1969).⁴

Berechnen wir die Bewertungsforschung aus einer *wissenschaftssoziologischen* Perspektive, läßt sich zweierlei festhalten: (1) Sie operiert mit einem klaren „Feindbild“ des Sozialen“ (Knorr-Cetina 1988: 85), demzufolge soziale Einflüsse verfallend auf die schulische Bewertung einwirken. (2) Sie löst Klausurarbeiten und Bewertungspraktiken aus ihren lokalen Kontexten heraus. Diese *Dekontextualisierung* stützt sich auf die Annahme, es handle sich bei der schulischen Prüfung um ein Meßverfahren, bei dem diejenigen, die messen, austauschbare Instrumente sind, die zu gleichen und somit gültigen Ergebnissen kommen (müssen), da sie lediglich die Ergebnisse *abbilden*. Die Beobachtung, daß die schulische Bewertungspraxis jedoch kaum den Gütekriterien experimenteller oder testpsychologischer Messungen genügt, führt dann zu einem Ruf nach objektiveren Verfahren (z. B. Schröder 1974; Gaude/Raschner 1971), nach „Instruktion“ und „Training“ der Lehrpersonen (z. B. Barnes 1985) oder nach Einführung von „Zeugnissen ohne Noten“ (z. B. Benner/Ramseger 1985; Lütger 1992).⁵

Schließlich ist die Bewertungspraxis noch Gegenstand der Schuldikdaktik. Sie unterscheidet die Bewertungsverfahren nach ihren Bezugsusername in „lenzielenorientierte“, „gruppenorientierte“ und „ispathe Norm“, auch „Ideal“, „Sozial-“ und „Individualnorm“ genannt. Bestimmen im ersten Verfahren definierte Lernziele die Bewertung, so orientieren sich Lehrpersonen in den beiden letztgenannten Verfahren an ihren Schülern – sei es als Kohorte, sei es als Einzelschüler (vgl. Schröder 1990; Heckhausen 1974). Die Frage, *wie* korrigiert und bewertet werden soll, wird zwar auch in der Schuldikdaktik (vgl. z. B. Langer et al. 1993; Bartnitzky 1989; Becker 1988; Baumann et al. 1979; Hurst et al. 1979; Schröder 1972), insbesondere

aber von *Bewertungspraktikern* selbst diskutiert.⁶ In praxisorientierter Art thematisiert: u.a. die Gestaltungsmöglichkeiten der Korrektur (z. B. Klemmisch/Osterhoff 1993; Schiffer 1991), Problemfälle (z. B. Althoff 1993; Speight 1991) sowie der Sinn der Korrektur überhaupt (z. B. Müller 1989; Stein 1989). Theoretisch, so läßt sich resümieren, blieb die Forschung über schulische Urteilspraktiken 'eingeklemmt' zwischen Reproduktionstheorie und Normativität.

Dieser Aufsatz nimmt nun gegen die Abstraktionen der Bewertungsforschung und gegen die korrespondenztheoretische Anlage der Bourdieuschen Praxistheorie eine *rekontextualisierende* Perspektive ein: Das direkte Bedingungsverhältnis von Makrostruktur und sozialem Handeln wird durch die Annahme, daß die soziale und symbolische Ordnung situativ verhandelt und hergestellt wird, ersetzt. Die Einbettung der Bewertungspraxis in lokale Kontexte basiert auf der Beobachtung, daß Lehrpersonen im Vollzug der Korrektur und Bewertung ihre Praxis der Wissensvermittlung und Wissensüberprüfung reflektieren. Ausgehend von diesem Phänomen der *Selbstbeobachtung* wird die Dominanzrelation verschoben: Anstelle der normativen Annahme der Bewertungsforschung, daß die Schülerleistung die Bewertung konstituieren soll (Dominanz der Sachdimension), wird die Konstruktionsleistung der Lehrpersonen fokussiert, in der sich die *Sozialdimension des Bewertens* entfaltet, ohne – wie in der Bourdieuschen „Maschine“ – auf die reproduktive Funktion sozialer Klassendifferenzen geeicht zu sein.

Was läßt sich nun unter der konstruktivistischen Prämisse, daß der Bewertungsprozeß die Leistung konstituiert, ethnographisch beobachten? Es läßt sich beobachten, daß die Bewertungspraktiker die Annahme hegen, ihren Bewertungen gingen immer schon Leistungen voraus. Hinter diese Annahme, die ihre Basis in der schulischen Temporalität von Lehren-lernen-prüfen-bewerten hat, ge-

6 Dieser Befund deckt sich mit Aussagen verschiedener Lehrpersonen, die mir im Laufe des Beobachtungsanfaltes gestanden, nicht während ihrer Ausbildung, sondern erst durch ihre berufliche Praxis gelernt zu haben, wie Klausuren zu korrigieren und zu bewerten sind. Die Beratungsliteratur beschränkt sich auf die Darstellung und Diskussion der „Ideal-“, „Sozial-“ und „Individualnorm“ (vgl. Gaude 1989), unterbreitet Vorschläge für Klausuren und Abiturprüfungen (z. B. Hinrichs/Milstein 1994; Schrems/Wolf 1985) und dokumentiert die Gesetzgebung (vgl. Rosenberger/Lindner 1984).

hen weder Bourdieu (Bourdieu/Saint Martin 1975; Bourdieu 1989: 48ff.) noch die Bewertungsforschung zurück. Entgegen dieser Fiktion,⁷ Leistung lediglich messend festzustellen, wird hier die Frage, ob eine Leistung an sich vorliegt, eingeklammert, denn sie steht dem Soziologen nicht als schon Gegebenes, sondern nur als Ergebnis eines Bewertungsvorganges zur Verfügung. Das heißt, daß der Beobachter soziale 'Verfälschungen' in der Leistungsbewertung nicht feststellen kann. Nicht die Qualität der Bewertung, nicht die (er-neute) Evaluierung des Lehrerteils und daraus resultierende Verbesserungsvorschläge, sondern die mit Bewertungspraktiken verknüpften Wissensprozesse sind hier von Interesse. Mit dieser ethnomethodologischen Indifferenz hält der Beobachter Distanz zu den von den Teilnehmern angenommenen Wirklichkeiten (vgl. Lynch 1993).

Methodisch wird zur Beschreibung der kulturellen Praxis des Humannessens auf die Ethnographie zurückgegriffen, wie sie in den Laborstudien zur Analyse naturwissenschaftlicher Wissensprozesse praktiziert worden ist (vgl. Latour/Woolgar 1979; Knorr-Cetina 1984; 1988; Lynch 1985; Amann 1990).⁸ Das empirische Material stammt aus einem neunmonatigen Beobachtungsaufenthalt in exklusiven Internatsschulen (vgl. Kalthoff 1994).⁹ Der Wunsch des teilnehmenden Beobachters bei der Korrekturarbeit von Klausuren oder der Notenführung

7 Fiktion wird hier im doppelten Sinne verstanden als Gemachtes, Erzeugtes sowie als Metapher für blinde Flecke (vgl. Knorr-Cetina 1994).
8 Zur Beschreibung angewandter Humanwissenschaften siehe Hirschauer (1993). Im Kontext der bildungssoziologischen und erziehungswissenschaftlichen Forschung scheint die bislang geübte Zurückhaltung gegenüber der ethnographischen Methode aufgegeben zu werden (vgl. beispielhaft Krappmann/Oswald 1995; Scholz 1994). Im Gegensatz dazu griff die Schulforschung in den USA schon früh auf die ethnographische Methode zurück (überblicksartig Farnen 1994; Terhart 1979).

9 Die Spezifik der untersuchten Internatsschulen wird hier nicht weiter berücksichtigt. Für die Gültigkeit des Datenmaterials ist dies aus zwei Gründen relativ unproblematisch: Zum einen beschränkte sich die Beobachtung der Klausurkorrektur auf Lehrpersonen, die ausschließlich als Lehrer im Schulbereich und nicht auch als Berater von Internatgruppen tätig waren. Das heißt, daß andere Aspekte der Sozialdimension nicht zur Geltung kamen. Zum anderen unterliegen die mündlichen Abiturprüfungen in privaten Internatsschulen den gleichen rechtlichen Regelungen, wie sie für das öffentliche Gymnasium gelten. Zur rechtlichen Kodifizierung des Prüfungswesens vgl. Kultusminister des Landes Nordrhein-Westfalen (1987); Holmeyer (1988); Niehus (1976).

dung anwesend zu sein, erzeugte bei den Lehrpersonen Ängste, Bedenken und Abwehr unterschiedlicher Art. Dennoch gelang es mir, bei fünf Lehrpersonen bis zum häuslichen Schreibtisch vorzudringen, sowie an acht mündlichen Abiturprüfungen inklusive Notenberatung teilzunehmen. Der Aufsatz will der Involvierung der Lehrpersonen nachgehen und fragen, wie sie ihre Rolle bei der Bewertung bestimmen, welche Fehlleistungen sie sich selbst zurechnen und welche nicht, wie stark oder schwach sie ihre Lernziele formulieren, und wie sie die Tatsache reflektieren, daß sie mit dem Effekt des eigenen Unterrichts konfrontiert werden, ja auf gewisse Weise in einen Spiegel schauen. Im folgenden wird die schulische Bewertungspraxis anhand der Korrektur von Klausuren als Verfahren implizierter Selbstbeobachtung beschrieben und analysiert (2.) und durch das kollektive Verfahren der Fremdbeobachtung kontrastiert (3.). Auf die soziale Situation der Prüfung, die von der Bewertungspraxis zu unterscheiden ist, wird hier nicht vertiefend eingegangen.

2. Die Klausur mit Selbstbeobachtung

Schule hat eine eigene Zeitform institutionalisiert: die periodische Zeit des Unterrichts (vgl. Schort 1990). Eine Basisannahme des schulischen Unterrichts unterstellt, daß Schüler kontinuierlich und unablässig mit neuem Wissen, das in der Schule die Form des „Stoffes“ annimmt, konfrontiert werden können, und daß sie – obwohl temporär ausgeschlossen – der Kommunikation „folgen“ (Kritisch hierzu Maekay 1975). Schüler bewegen sich in diesem sich anbahnenden Material: ihre Antworten zeigen den Lehrpersonen, „wo die Schüler sind“ (Lehrer). Die Form, in der „Stoff“ bearbeitet wird, kann als *symptomatische Wissensüberprüfung* bezeichnet werden. Das, was Lehrpersonen *sehen* können, ist abhängig von einer Sortierung der Schüler entlang von Frageformen.¹⁰

In die periodische Zeit des Unterrichts hat die Schulorganisation Momente der *systematischen Wissensüberprüfung* installiert: die schriftliche

¹⁰ Die Erzeugung von Wissen im mündlichen Unterricht etabliert unterschiedliche Wissensstandards durch Fragestellungen, die sich an Leistungsniveau der Schüler orientieren. Diese Aktivität unterliegt dem Diktum der Unsichtbarkeit: Die Vorsortierung der Fragen entsprechend der Norenskalisierung der Schüler darf nicht offensichtlich werden (vgl. hierzu ausführlicher Kalthoff 1995), zur Bewertungssystematik des mündlichen Unterrichts vgl. Streeck (1979).

1. Ausschüttung von Wissen: Stunde um Stunde
2. Fragetyp und Verteilung der Schüler im Wissenskörper (symptomatische Wissensüberprüfung)
3. Fixierung der Schüler im Bewertungsverfahren (systematische Wissensüberprüfung)

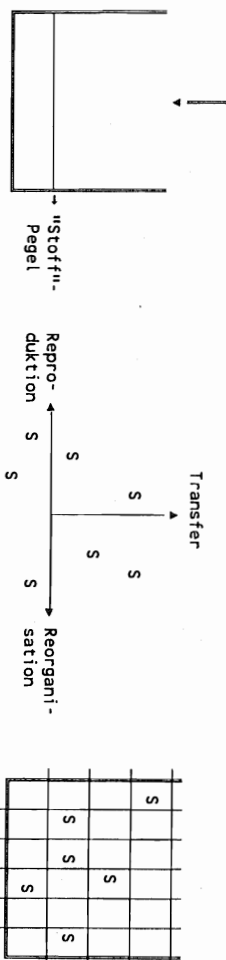


Abb. 1 Das Modell des Wissenscontainers.

Prüfung. Diese – so wird angenommen – fixiert die Plätze der Schüler im Wissensmaterial und somit in der Klassifikation.¹¹

Der Zeitpunkt der Fixierung wird durch die Zerfallszeit des Schülerwissens bestimmt, also durch die Notwendigkeit, Wissen zu einem Zeitpunkt abzufragen, wenn es noch abfragbar ist. Ein Lehrer:¹²

„Also wer in die zweite Klasse eintritt mit Geschichte und man glaubt, man könne da was aus der zehnten machen, da is NICHTS mehr da, ja, nichts. Also in zwei Jahren is alles vergessen. Was geblieben is, is vielleicht n Verständnis für bestimmte Zusammenhänge, wenn se n guten Unterricht gehabt ham. Sie wissen irgendwie n bißchen was, so groß können se schon sagen, daß der zweite Weltkrieg fünfvierzig aus war, aber so als einzelne Fakten, es is nichts da.... Die haben aber auch nicht die Ruhe und die Zeit, die Sachen zu vertiefen. Und so wird dann eine Arbeit nach der andern geschrieben, und dann wird ein Ding nach dem anderen auswendig gelernt, im Hirn abgeakt, und Platz geschaffen fürs nächste Fach.“

Der periodische Zeitbegriff der Schule konstituiert eine ebenso strukturierte Wissensaneignung, auf die die schulische Organisation mit einem Rhythmus von mündlicher Wissensproduktion und -reproduktion im Unterricht und ihrer schriftlichen Überprüfung in der Klausur antwortet.

¹¹ Die schulische Wissensproduktion und Wissensüberprüfung als Wissenscontainer zu denken, in dem Wissen nach spezifischen Regeln organisiert und präsentiert wird, geht auf eine Anregung von Klaus Amanan zurück (vgl. auch Amanan 1990: 7ff.). Der Buchstabe S steht hier für einzelne Schüler.

¹² Als Transkriptionszeichen werden folgende Zeichen verwendet: SCHULE = laute Aussprache; Schule = betonte Aussprache; ... = Auslassung in der Transkription; ... = Dehnung; ? = steigende Betonung; – = Abbruch einer Äußerung.

2.1 Das Bewertungssystem

In der Klausurzeit fabrizieren Schüler einen Text, der zwischen Schüler und Lehrperson zirkuliert und über den letztere ein Urteil fällen.¹³ Für Lehrpersonen ist die Korrektur nicht allein ein Abhaken 'richtiger' oder 'falscher' Schülerantworten, d. h. Beobachtung einer fremden Leistung, sondern ein ambivalenter Moment, der unter Umständen eine zentrale Prämisse gefährdet, nämlich die Annahme, daß sie Schüler und „Stoff“ genau kennen. Die Unsicherheit des Ausgangs – es kann die Voraussetzung für Zurechnungsalternativen und für das Phänomen der Selbstbeobachtung.

Der Ambivalenz begegnen Lehrpersonen damit, daß sie die Korrektur hinauszuögern und in der Regel an ihren hässlichen Schreibstil verlagern, der die Korrekturarbeit vor äußeren Blicken abschwächt. Lehrpersonen gehen mit den Schülertexten gewissermaßen in Klausur, deren Bedingung die Abdunkelung ist.

Zugleich hegen Lehrpersonen Zweifel daran, ob das, was sie da lesen, Schülerwissen abbildet; es ist ein ganz grundlegender Zweifel an der Prüfung zugeschränkter Funktion der Leistungsbeurteilung. In diesen Fällen reflektieren sie nicht auf sich, sondern auf die Abbildungsfunktion der Klausur. Eine richtige Schülerantwort bedeutet für sie nicht zwingend, daß der Schüler „es verstanden“ hat. Ein Lehrer:

¹³ Am Ende einer Klausur wissen Schüler zwar, was sie geschrieben haben (sie kennen Quantität und Inhalt ihres Textes), sie wissen aber nicht, was sie (geschrieben) haben; dies wissen sie erst, wenn sie den zensurierten Text zurück bekommen.

„Wenn se nen Kreislauf (des Land-See- und Systems, H.K.) zeichnen, zeigen se eigentlich, daß se die Dinge schon wohl wissen müßten – eigentlich. Vielleicht ham se auch nur auswendig gelernt, das kann man net sagen, aber sie ham nimmer explizit formuliert. Den Kreislauf kann ich auswendig lernen, ohne ihn zu verstehen, ja? Viele Schüler machen das. Aber das kann ich net mehr nachprüfen.“

Die Unmöglichkeit, das dokumentierte Wissen daraufhin zu überprüfen, ob es verstanden oder auswendig gelernt wurde, läßt Lehrpersonen nur eine Wahl: Sie geben den Punkt.

In der praktischen Vorgehensweise lassen sich die Bewertungssysteme danach unterscheiden, wie Lehrpersonen ihre Korrektur und Bewertung organisieren. Es konnten Lehrpersonen beobachtet werden, die das „kriteriums-/lernzielorientierte“ Verfahren anwendeten: Sie schafften sich vor der Korrektur eine Musterlösung, die sie „Erwartungshorizont“ nennen. Dieser Horizont kombiniert die Wissens Elemente, die Lehrpersonen in ihren Fragen verstecken, mit Bewertungen, deren Gewicht in Zahlen („Punkte“) ausgedrückt wird. Die Höhe der Zahl gibt an, ob die Beantwortung der Frage als schwer oder leicht, als aufwendig oder reproduktiv, als bedeutsam oder nicht bedeutsam eingestuft wird. Diese Soll-Ausstellung macht aus Wissen Zahlen, mit denen sich rechnen und ökonomisch kalkulieren läßt. Schriftlich fixiertes Wissen der Schüler wird bilanzierbar. Die Korrektur mit dem „Erwartungshorizont“ funktioniert wie ein Abgleich zwischen der Sollseite des Wissens und den Istzuständen in den Schülertexten. Der „Erwartungshorizont“ ist ein reflexives Instrument; er demonstriert den Lehrpersonen ein vom Schülerwissen unabhängiges Anspruchsniveau.

Vom Kalkulationstraster der Punkteverteilung ist eine Liste zu unterscheiden, die Lehrpersonen ad hoc zu Beginn der Korrektur als eine mitunter vage Vergegenwärtigung ihrer Erwartung formulieren. So fragte sich zum Beispiel ein Lehrer, der das „gruppenorientierte“ Verfahren anwandte, zu Beginn einer Korrektursitzung, was er denn alles hören wolle und kritzelte verschiedene Antworten auf einen Zettel, ohne diesen Punktwerte zuzurechnen. Ändere Lehrpersonen verzeichneten ganz auf diese Form der Vergegenwärtigung von Erwartungen und konzentrierten sich allein auf das in den Klausurarbeiten fixierte Wissen der Schüler. Oft sind sie es, die ihre Zensurvorstellung entsprechend den Schülerleistungen definieren, wobei das Feld des Machbaren insbesondere durch die besseren Schüler abgesteckt wird.¹⁴ Manche von ihnen

systematisieren das Verfahren, indem sie ihre Schüler vor der Korrektur in Leistungstapel sortieren: oberer, mittlerer, unterer Leistungsbereich. Die Vorsortierung kann auch nach inhaltlichen Aspekten erfolgen, und zwar dann, wenn in der Klausur verschiedene Themen angeboten werden. So im Fach Deutsch einer 13. Klasse, in der ein Lehrer eine Kurzgeschichte, einen Auszug aus einem besprochenen Roman, einen freien Text und ein Gedicht zur Auswahl stellte. Die unterschiedbaren Schwierigkeitsniveaus ergeben eine „Auswahl im Vorfeld“, so der Lehrer. Die in der Korrektursitzung nebeneinander liegenden Stapel der Schülertexte stehen dann nicht allein für eine thematische Sortierung, sondern ebenfalls für eine nach dem Leistungsvermögen. Fixiert der „Erwartungshorizont“ einen Maßstab entlang der gestellten Fragen vor der Korrektur, so etabliert die vorsortierende Form ihren Maßstab nach der Korrektur der Arbeiten vermeintlich besser Schüler. Die Muster unterscheiden sich in der Konstitution ihres Bezugssystems und in der Wahl des zeitlichen Moments, der vorangegangene Schülerleistungen in das Verfahren einbindet. Kurz, die Bewertungsmuster stehen für unterscheidbare Selektionsmodi, die aber den Zweck (Klassifikation) unangestast lassen.

2.2 Der Verteilungsvorgang

Vor der ersten Lektüre schauen sich Lehrpersonen das „Material“ an; die Quantität der Texte wird festgesetzt oder abgeschätzt, einige Texte werden durchgeblättert. In der ersten Korrekturlektüre, mit der sich Lehrpersonen einen Überblick verschaffen wollen, annotieren sie orthographische, grammatikalische und inhaltliche Fehler mit einem Rotstift bzw. mit Bleistift. Punktzahlen werden bei jeder Aufgabe mit Bleistift am Rand, auf Zetteln oder in einer Kreuztabelle notiert; erste Bemerkungen für das abschließende Urteil werden auf Zetteln festgehalten, wie z. B. „Erfüllte den des Schemas fehlt völlig“. Verschiedene Kürzel werden an den Rand geschrieben, wie z. B. „Gr“ (Grammatik), „SF“ (sachlicher Fehler), „SM“ (sachlicher Mangel), „L“ (Fehler in der Logik), „+“ (gut, richtig), „?“ (unklar) sowie Auslassungszeichen, Unterstrichungen und Bemerkungen (z. B. „Lesbarkeit“, „Argumentation?“). Bei den Lehrpersonen, die mit einer Musterlösung arbeiten, wandert der Blick in regelmäßigen Abständen vom Schülertext zur Musterlösung, so als wollten sie den Ausschlag des Pegels ablesen.

¹⁴ Für Luhmann/Schorr (1988: 241 f.) ist die Orientierung der Lehrperson an den „besten Schülern“ eine historisch tradierte Methode der Selektion.

Während der z.T. sehr langen Korrektursitzungen stehen Lehrpersonen über den Text in einem (fluktuiven) dialogischen Verhältnis mit den einzelnen Schülern: „So lieber Uwe, was mach ich jetzt mit dir?“ Die Texte produzieren Gefühle wie Freude über eine gelungene Arbeit, Überraschung über ein nicht erwartetes positives Ergebnis, Enttäuschung über eine mifflarene Klausur. Lehrpersonen können buchstäblich leiden. Sie setzen und konstatieren mit einem distanziererten Blick Unverständnis über das, was sie lesen: „Was schreibt die denn für ein Zeug“ oder „Der hat die ganze Aufgabe nicht verstanden“.

Eine Lehrerin korrigiert eine Klausur: die sie später mit „fünf“ bewerten wird. Mit der linken Hand stützt sie ihren Kopf ab, ihr Blick signalisiert Skepsis. Sie beugt sich nah über den Text, so als könne sie nicht glauben, was sie da liest. Laute wie „tse, tse, tse“ entfahren ihrem Mund. Kurz darauf streicht sie mit einer energischen Geste mehrere Zeilen mit einem Federstrich durch. Die Hand, die ihren Kopf stützt, vermag ihn nicht mehr zu halten; er rutscht ab, so als sei die Arbeit nicht mehr zu retten. Sie kommentiert dem Beobachter ihre Empfindungen: „Wenn Sie jetzt nicht da gewesen wären, hätte ich laut Scheiß geschrien“.

Lehrpersonen haben Schüler vor ihrem „geistigen Auge“ und hegen Erwartungen an den Text des Schülers; sie erkennen seinen Stil, seine „Schreibweise“: „Man *kenn*t auch die Art, die kennt man. Dann denkt man, ja jetzt kommt se ja bald zum Hauptteil, und dann freut man sich, AH JA, jetzt müßte doch *des* kommen, und ach da kommts, ach warum kommts nich. Also das geht schon so. Man kennt das.“ Lehrpersonen unterscheiden die Klausurtexte neben der Adäquatheit auch nach ästhetischen Merkmalen: „Es ist schön sie zu lesen“ und „Auch ein schöner Schreiber“.

Lehrpersonen verändern die Leseform der Korrektur, wenn sie von einer gelungnen zu einer mifflungnen Arbeit wechseln (oder umgekehrt). Während die Korrektur einer guten Arbeit einer konzentrierten Lektüre gleicht – Bleistift oder Fingergelb dem Text –, gehen Lektüre und Rosistift-Korrektur einer schlechteren Arbeit ineinander über. Mit einer guten Arbeit, zu der Lehrpersonen eine andere libidinöse Beziehung zu unterhalten scheinen, gehen sie liebevoll um, mit der schlechten rigoros. In ihren Gefühlen drückt sich die Art bei der Zurechnung aus. Der Affekt ist eine Zurückweisungs- oder Anerkennungreaktion, mit dem Lehrpersonen ihre Identifizierung oder Abwertung ausdrücken. Im negativen Fall bewerten Lehrpersonen nicht, sondern sie werten ab.

Lehrpersonen entziffern ihren Eigenanteil an mangelnden Schülerleistungen. Ob diese sie etwas

„angehen“ oder nicht, unterscheiden sie auf Grundlage ihres intimen Wissens über den „Stoff“ und die Klasse. Hierzu verschiedene Lehrpersonen:

„Du überlegst auch, woran liegst, da machste dir schon Gedanken. Warum ist es so schlecht ausgefallen? Vielleicht war einfach mein Erwartungshorizont zu hoch.“

„Bei'n zwölften hatten wir'n neues Buch, das war vor'm Jahr eingeführt worden. Das Buch is viel zu anspruchsvoll. Und bei der ersten Klausur hab ich denen viel aus'm Buch aufgegeben. Und das hat die total verwirrt. Also da war für mich klar, warum die so schlecht ausgefallen war.“

„Bei der letzten Arbeit hab ich einfach gesehn, daß die selbst die einfachsten Aufgaben, die nur Reproduktion waren, einfach nich gelernt hatten. Die hatten *rix* gemacht. Und dann sag ich mir, stopp, halt mal, dann is das kein Punkt, der mich was angeht, ja? Dann is das ihr Bier.“

Lehrpersonen suchen nach Alternativen, denen schlechte Ergebnisse zugeschrieben werden können; es kann ein schwieriges Buch oder das Thema selbst sein, die unzureichende Erklärung im Unterricht oder die zu hohe Erwartung. Auf die mangelnde Vorbereitung der Schüler – Lehrpersonen sprechen von fehlender „Leistungsreife“ – folgt eine eindeutige Schulzuweisung. In diesen Fällen strafen Lehrpersonen – auch aus Enttäuschung über einzelne Schüler – mit schlechten Noten.

Lehrpersonen wissen auch, bei welchen Themen ihres Faches (z. B. bestimmte Gebiete in der Chemie) die Klausuren regelmäßig schlecht ausfallen. In diesen Fällen wird das Ergebnis dem Thema zugeschrieben: mit dem Themenwechsel im Unterricht „gleichlich es wieder aus“, so ein Lehrer.

Die Selbstbeobachtung bezieht sich ebenfalls auf die wiederholte Lektüre der Klausurarbeiten oder einzelner Aufgaben. Zwei Lehrpersonen:

„Du guckst dann immer wieder. Da hat er vielleicht noch was richtig. Also was er schreibt is *völlig* falsch oder *ziemlich* falsch, aber du akzeptierst es doch.“

„Ja gut, *dar* is der Zusammenhang. Ich weiß, sie meints eigentlich, es is ober noch die falsche Fachsprache. Da kann ich aber nochmal n halben Punkt geben. Das kann ich machen.“

Die intensive Lektüre stellt eine Interpretationsressource dar, mit der Lehrpersonen einzelnen Schülerantworten richtige Antwortelemente entlocken können. Lehrpersonen können aus dem, was nicht bewertbar ist, etwas machen, das als bewertbar erscheint; sie verschieben die Differenz von „noch bewertbar“ und „nicht mehr bewertbar“. Ob Lehrpersonen etwas „rausohlen“ können oder nicht, hängt davon ab, ob sie noch Bewertbares „finden“, das sich in Punkten ausdrücken läßt. Es ist fast so, als unterschieden Lehrpersonen im

Zweifelsfälle für ihre Schüler, denn sie lassen eher falsche als richtige Antworten *durchgehen* und zeigen sich wohlwollend gegenüber dem Klausurtext. Wohlwollen wird ebenfalls als Anpassung an die Leistung der Schüler in der Differenz zwischen erster und zweiter Korrektur sichtbar. Ein Lehrer:

„Dann hab ich auch schon beobachtet, daß ich dann wesentlich strenger korrigiere am Anfang, im ersten Durchgang viel schlechtere Noten erstmal gebe. Ich weiß nich warum, also pu: : . . . vielleicht is man erstmal enttäuscht oder so, ja?, man hat ja n Erwartungshorizont vor sich. Und später gleichlich das wieder n blöchen an.“

Dem Wissen um diese Differenz entspricht die Praxis der Korrektur: Zunächst halten Lehrpersonen ihre inhaltlichen Urteile reversibel; der Gebrauch von Bleistift, Radiergummi und Notizblättern zeigt ihren Vorbehalt und sichert die Möglichkeit, erste Urteile problemlos zu revidieren.

Am Ende des ersten Durchganges wird die Gesamtpunktzahl in einer Kreuztabelle notiert und/oder ein erster Kommentar mit Notenbereich auf einem Notizzettel geschrieben. In der Kreuztabelle entsteht ein nach Antwortelementen, Aufgaben und Gesamtpunktzahl sortiertes Schülerarray. Vorläufige Urteilsbegründungen lauten z. B. so: „Der Martin kriegt ne eins. Auch wenn er nicht im Einserbereich liegen sollte, kriegt er sie. Dann ist es eben eine schwache eins.“ Die zweite Korrektur erfolgt als Kontrolle der ersten. Lehrpersonen kontrollieren, ob sie richtige Antwortelemente übersehen und die richtige Anzahl der Verrechnungspunkte notiert haben; sie verwenden Bleistiftziffern in rote Tinte-Ziffern. Dieser Vorgang sei, so eine Lehrerin, bei den Arbeiten am intensivsten, die „auf der Kippe stehen.“ Die Gesamtpunktzahl wird festgelegt und mit der ersten verglichen; ein abschließender Kommentar wird vorformuliert oder direkt am Ende des Schülertextes notiert. Mit der wiederholten Textlektüre sichern sich Lehrpersonen gegen Fehler ab und bestätigen (oder korrigieren) die Verteilungsstruktur. Sie machen ihre Bewertung dicht.

2.3 Die Arbeit am Schnitt

Von der Feinjustierung einzelner Klausurtexte ist die Arbeit am „Schnitt“ zu unterscheiden. Stellen Lehrpersonen im Korrekturvorgang fest, daß eine Klausur „abzustützen“ droht oder eine Aufgabe vom Großteil der Schüler nicht beantwortet werden konnte, redefinieren sie ihren Maßstab, denn jetzt geraten nicht nur die Schüler, sondern auch sie unter Druck. So meinte ein Lehrer während der Korrektur:

„Und zur Gewichtung der Klausur sag ich jetzt auch schon. Sollte der erste Teil wesentlich besser gelöst sein von den Schülern, und der zweite ganz marode und ich *drohe* mit der Hälfte der Leute unter fünf zu kommen, dann erlaube ich mir nachträglich, meine Gewichtungen, meine Kriterien zu verschieben.“

Der Bedarf zur Korrektur entsteht in dem Moment, wo das Ergebnis nicht mehr „paßt“. Unabhängig vom Bewertungsmuster fungieren insbesondere gute Schüler als Alarmzeichen, an denen sich Lehrpersonen orientieren. Macht die erste Korrekturkritik ein schlechtes Ergebnis wahrnehmlich, verschieben Lehrpersonen „Gewichtungen“ und arbeiten so am Schnitt der Klausur. Mit ihm verändern sie nicht die Verteilungsstruktur der Schüler, sondern verschieben die Klasse insgesamt in höhere Notenbereiche. Dieser Vorgang soll anhand von zwei Beispielen dargestellt werden.

(1) In seiner Bewertung geht der Lehrer nach dem „kriterienorientierten“ Verfahren vor. Er erstellt einen „Erwartungshorizont“, mittels dessen er 24 Verrechnungspunkte auf die fünf Klausurfragen verteilt. Nach der ersten Korrektur von acht Schülern, deren erreichte Punktzahlen er für die einzelnen Aufgabenteile in eine Tabelle einträgt, stellt er fest, daß die Antworten auf eine Frage stark von seinem Erwartungshorizont abweichen: „Nach den ersten Arbeiten war klar, die können scheinbar nicht alle fünf Punkte kriegen. Die hat keiner.“ Einen wichtigen Anstoß, die Musterlösung an die Schülerleistung anzupassen, liefert der Text einer „sehr guten Schülerin“, die die Aufgabe auch nur unbefriedigend beantwortet konnte. In den Momenten also, in dem der Lehrer, der sich zunächst an seinen Lernzielen orientiert, mit dem Ausgang der Klausur im Abwind ist, greift er auf die historisch alte Selektionsmethode der Orientierung an den „besten Schülern“ zurück. Er begründet es so: „Ich hab das jetzt gemacht, weil die Arbeit war einfach vom Schnitt her nicht adäquat von dem was ich eigentlich erwartet hatte.“ In der Ursachenforschung formuliert er Unsicherheit, rechnet sich aber einen erheblichen Anteil selbst zu: „Es zeigt, daß ich es entweder falsch eingeschätzt hab, oder daß es net richtig gut erklärt war“.

In seinem „Erwartungshorizont“ hatte der Lehrer für die einzelnen Antwortelemente der Frage eine aufgebildete Punkteverteilung vorgesehen. In der Überarbeitung der Punkteverteilung für diese Aufgabe streicht er die beiden Teilantworten heraus, die die Schüler bisher nicht gebracht haben. Daraus resultiert eine Verschiebung der Schüler in

höhere Notenbereiche. Die Möglichkeit der nachträglichen Korrektur dokumentiert den Gestaltungsspielraum im Bewertungsprozess – und auf diese Weise die Konstitution von Ergebnissen, die „adäquat“ sind. Im Vollzug der zweiten Korrektur wird das Verhältnis von Lehrer („ich“) und Schülerleistung angepaßt.

(2) Im zweiten Beispiel arbeitet der Lehrer nach dem „gruppenorientierten“ Verfahren. Zunächst stellt er die Textquantität fest (elf Schüler produzierten 34 Seiten Text), um sich anschließend das Thema der Klausur nochmal zu vergegenwärtigen. In Stichworten notiert er einige der in den Fragen versteckten Themen. Die beiden Klausuraufgaben will der Lehrer mit ein Drittel (erste Aufgabe) und zwei Drittel (zweite Aufgabe) gewichten. Nach der Sortierung der Klausurtexte in Notenbereiche („handliche Happen“) – oberer Notenbereich vier Klausurtexte, mittlerer Bereich drei Texte, unterer Bereich vier Texte – macht er sich an die Korrektur der Texte des oberen Notenbereiches, aus dem er die Kriterien seiner Verteilung gewinnen will. Diese Texte sollen als Orientierungspunkte für die Streuung der anderen Schüler fungieren: „Ich möchte mich an den Guten orientieren, das ist eine Rückkoppelung“. Die Korrektur des vom Klassenprimus erstellten Klausurtextes fällt aber dann nicht so aus, wie erwartet. Den ersten Teil würde der Lehrer mit „drei minus“ oder „vier plus“ bewerten. Damit der Schüler aber noch auf eine entsprechende Note kommt – „Ich kann dem keine vier geben“ – wird der erste Teil in einer zweiten Lektüre auf „drei“ gesetzt.¹⁵ Der zweite Teil „muß zwei sein“, damit der Lehrer in der Gesamtnote eine „zwei minus“ oder „drei plus“ begründen kann. Nach der Lektüre der vier Texte des oberen Notenbereiches, d.h. in Relation zu anderen, vorselektierten Klausurtexten, plaziert der Lehrer die Arbeit mit einer „drei plus“. Am Ende des ersten Korrekturdurchganges verfügt er über einen Eindruck der Texte. Seine ursprüngliche Gewichtung der beiden Klausurteile verschiebt er jetzt von ein Drittel/zwei Drittel auf eine Gleichgewichtung beider Teile. Das heißt, derjenige Klausurteil wird stärker gewichtet, in dem die Schüler des oberen Notenbereiches die besseren Leistungen erbracht haben. An diesem Beispiel wird nicht allein die Orientierung an der in der Klausur erbrachten

Schülerleistung deutlich, sondern insbesondere die Orientierung an den bislang etablierten Verteilungen und erbrachten Leistungen. Der Lehrer stabilisiert die Rangfolge der Schüler, indem er bisherige Ergebnisse als aktuellen Verteilungsschlüssel einsetzt, der eine bestimmte Verteilung wiederholt; Resultat des Vorganges sind reduzierte Mobilitätschancen der Schüler auf der Notenskala.

Abgeschlossen wird der Prozeß durch den „Schritt“. Als Mittelwert¹⁶ sagt er wenig über die reale Verteilung der Noten; er gibt die Tendenz der Verteilung an. Der „Schritt“ ist die fiktive Note des Klassenkörpers in der rationalisierten Form der „Rückkoppelung“, er suggeriert eine gemeinsame Note der Schülerkohorte, so als habe diese eine 3,2. Mit dem Mittelwert wird der Klausurzyklus beendet, der die Verteilung der Schüler des Wissensstransfers, den die Lehrperson organisierte. In diesem Sinne reflektiert der Schnitt ihre Selbstbewertung.

2.4 Zwischenfazit

Den individuellen Schülertext entlang einer Punkte- bzw. Notenskala als gut oder schlecht definierend, bringen Lehrpersonen den Schülertext in Relation zu anderen Schülertexten. Lehrpersonen definieren die Abstände der Schüler; sie ordnen Schüler, indem sie sie auf Notenziffern setzen; auf diese Weise nehmen Schüler einen Platz in der Klasse und einen Platz in der Klassifikation ein. Dabei erfüllt die Klausur in der Funktion der Differenzierung nur dann, wenn die Werte differieren. Für Lehrpersonen ist es undenkbar, alle Schüler auf nur ein oder zwei Rangplätze anzusetzen; es gibt in ihrer Vorstellung einfach gute und weniger gute Schüler. Das heißt auch, daß die Feststellung der Schülerleistung von der Funktion, Schüler

über eine Skala zu verteilen, dominiert wird.¹⁷ Über den Aufgabenkatalog und über ihr Bewertungsmuster organisieren Lehrpersonen die „Auftrennung“ (Lehrer) der Schülerkohorte. Die Schüler sollen sich über die Notenskala *verteilen*; dies ist die Normalität der Verteilungsarbeit. Für Schüler heißt dies, es können nie alle „sehr gut“ werden. Nicht die Gleichheit ist das Ziel, sondern die Differenz, die ihrerseits soziale Geltung beansprucht. In der Etablierung der Differenz materialisiert sich die soziale Selektionsfunktion der Schule.

Der Verlauf der Klausur kommt einem mehrstufigen Reduktionsprozeß gleich. Die Intensität des Erarbeitungsprozesses und die Atmosphäre im Unterricht, die Selbstreflexion der Lehrpersonen über ihre Unterrichtstätigkeit sowie die Anforderungen des Lehrplanes bestimmen den Kontext, durch den die Klausur geschleust wird. Der Verlauf ist an mehreren Stellen offen gehalten für Anpassungen an die Norm, der sich Lehrpersonen verpflichtet fühlen. Kriterien- oder gruppenorientierte Bewertungssysteme erweisen sich in ihren Händen als flexible Instrumente, die zwischen Lernzielen und Schülerleistung austariert werden können.

3. Die Klausur mit Fremdbeobachtung

Wenn Lehrpersonen schriftliche Klausuren bewerten, tun sie es – wie gezeigt – allein. Die Konstellation sieht auf der einen Seite eine Lehrperson vor, auf der anderen Seite eine definierte Anzahl von Schülern. Daß, wie behauptet, Lehrpersonen mit ihrer Bewertungsarbeit der Schülerklausur immer auch ihre eigene Leistung konstruieren, wird dann offensichtlich, wenn die Schulorganisation diese Konstellation aufhebt. Dies ist zum Beispiel bei der Abiturprüfung der Fall.¹⁸ Im schriftlichen Teil der Abiturprüfung erfolgt die Involvierung eines zweiten unter Umständen auch dritten Korrektors als ein Nacheinander: Der zweite bzw. dritte Korrektor weiß immer schon, wie der Klausurtext vom ersten bzw. zweiten Korrektor bewertet worden

ist.¹⁹ Anders in der mündlichen Abiturprüfung: Hier werden die Noten in einem Gremium von drei oder vier Lehrpersonen ausgeteilt. Ich werde mich im folgenden auf die mündliche Abiturprüfung konzentrieren, um an diesem Beispiel die Praxis der kollegialen Prüfung als Fremdbeobachtung zu beschreiben.

3.1 Das Prüfungsgespräch

An den Tagen der mündlichen Abiturprüfung dominiert eine fein ausgetiffelte Organisation die Bewegung der Prüflinge. Einem zeitlichen Schachbrettmuster vergleichbar, werden die Prüflinge, die zu bestimmten Zeiten an bestimmten Orten vorgelesen sind, in Etappen von peripheren Räumen zum Zentrum – dem Prüfungsraum – geleitet. Sinn und Zweck des Prozederes ist der individualisierte Prüfling und somit die Wahrung des Prüfungsgeheimnisses, d.h. die Verhinderung illegitimer Kontakte und Vorbereitungsmöglichkeiten. Zugleich markiert dieses Verfahren für den Prüfling die Bedeutung des Rituals, zu dessen Objekt er wird. Die Prüflinge lassen sich nach den Gründen, die sie in die mündliche Abiturprüfung treiben, unterscheiden in reguläre und zusätzliche Prüflinge; letztere differenzieren sich nochmals in Abweichter und Gefährdete. Die Regulären werden programmgemäß im vierten Abiturfach mündlich geprüft, die zusätzlichen hingegen treten auf Problemfälle hin. Solche Problemfälle sind

¹⁵ Dies kann für Lehrpersonen zu überraschenden und unangenehmen Ergebnissen führen. In einer beobachteten Schule wurden einem Lehrer gleich mehrere Schüler, deren schriftliche Abiturarbeiten er im „Dreierbereich angesiedelt“ hatte, durch den anonymen Zweit- und Drittkorrektor in den Bereich der Note „Fünf“ herabgesetzt. Entsprechend die Benotung durch die Lehrperson der Note der Schüler; so resultierte aus der endgültigen Note eine „Abweichung“. In der „Abweichung“, d.h. in der Diskrepanz zwischen Schülerbewertung im Unterricht und im Abitur, drückt sich das Abschneiden der Lehrperson aus. Zur Bedeutung der Abiturprüfung für die berufliche Qualifikation meinte ein Lehrer: „Man darf nicht vergessen, es ist eine Prüfung vor der Schulleitung. Das ist ganz wichtig. Ich bin überzeugt davon, daß ein Schulleiter sehr genau sich anschaut, was für Listen da rein und raus gehen an Noten und wie korrigiert wurde, die Ergebnisse. Und daraus ergibt sich natürlich auch die Beurteilung für einen selbst im Beruf als Lehrer. Und in Beurteilung, die man bekommt, kann man das auch lesen zum Teil, daß es also heißt, es wurden so und so viele Abiture abgenommen, mit guten Ergebnissen, mit hervorragenden Ergebnissen, die auch dem entsprechen, was unterrichtet wurde.“

¹⁶ In den Schulen ist die Errechnung des Mittelwertes gängige Praxis; Lehrpersonen (oder auch der Zentralcomputer einer Schule) verrechnen die Noten. Diesem Vorgang unterliegen zwei notwendige Schritte. Eine definierte und linear geordnete Menge von Prädikaten („sehr gut“, „gut“, ...) wird in Form ebenfalls linear geordneter Ziffern („1“, „2“, ...) ausgedrückt (Ziffernoten). Bei der Mittelwertberechnung betrachten bewertende Lehrpersonen die Ziffern als Zahlen, mit denen sich rechnen läßt. Dabei würde der Charakter der Notenskala als Ordinal- oder Rangskala gerade dieses statistische Verfahren der Bildung von Mittelwerten unzulässig machen (vgl. Fischer 1991; Orth 1974).

¹⁷ Lehrpersonen, die nur einen kleinen Teil der Skala ausnutzen würden, handeln sich bei Lehrern und Schülern den Ruf ein, „weich“ oder „zu weich“ zu bewerten.

¹⁸ Die Abiturprüfung ist nur eine Form der institutionellen Fremdbeobachtung. Im Rahmen ihrer beruflichen Laufbahn werden Lehrpersonen in bestimmten zeitlichen Abständen von Schülern beobachtet und beurteilt (vgl. Hopf 1983).

positive oder negative „Abweichungen“ sowie nicht ausreichende Punktzahlen in den schriftlichen Prüfungen, die das Bestehen des Abiturs gefährden (vgl. Böhm et al. 1994: 188 ff.).

In den Prüfungen verteilen sich die Prüfer²⁰ im Prüfungsraum, wobei sie eine unsichtbare, aber deutliche Trennungslinie zum Prüfling ziehen. Sie achten auch darauf, daß der Prüfling (d. i. der Fachlehrer) dem Prüfling am nächsten sitzt, daß das Kommunikationspaar Prüfler-Prüfling den Ablauf gestaltet. Nach einer kurzen Einweisung beginnt der Prüfling mit der Lösungspräsentation, deren Dauer in den beobachteten Prüfungen von zwei bis zu dreizehn Minuten variierte. In den mathematisch-naturwissenschaftlichen Fächern werden von den Prüflingen Versuchsabauten erklärt oder Aufgaben gerechnet, wozu sie häufig die Tafelschrift als Visualisierungsmittel einsetzen; in den geisteswissenschaftlichen oder Sprachfächern sind i. d. R. Texte zu interpretieren. Die Prüfler greifen in den Vortrag nur dann ein, wenn der Prüfling stockt oder sich verirrt hat; in diesen Augenblicken geben sie kleinere Hinweise, damit der Prüfling seinen Vortrag fortsetzen kann. Lehrpersonen als Prüfler tendieren dazu, mit ihren Interventionen so lange zu warten, bis der Prüfling das Ende seiner Ausführungen markiert hat („Das wars“). Dieser Ablauf des ersten Teils der mündlichen Abiturprüfung kehrt die unterrichtliche Situation: Nicht mehr die Lehrperson hat das Recht des ersten Wortes, sondern der Prüfling. Die Abgabe des Rederechtes an den Prüfling wird durch eine Reihe von Vorsequenzen der Prüfung eingeleitet: durch das kontrollierte Begleiten von einem Raum zum anderen, durch die Vorbereitungszeit und durch die Übergabe des Rederechtes durch den Vorsitzenden in der Prüfung selbst. Das Rederecht bedeutet für den Prüfling aber auch die Verpflichtung zu sprechen: er muß nicht mehr – wie im

Unterricht – zuhören, er muß sprechen: Er ist *dram*.²¹

Den Ausführungen des Prüflings schließt sich ein Frage-Antwort Spiel zwischen Prüfler und Prüfling an; nur selten ergreifen die anderen Prüfler das Wort. Einem schweigenden Publikum vergleichbar sind diese Lehrpersonen partiell anonymisiert. Je nach Qualität der Ausführungen des Prüflings formuliert der Prüfler seine Anschlussfragen. Es wurden Lehrpersonen beobachtet, die sich absteinert gegenüber den Antworten und Kommentaren ihrer Prüflinge verhalten. Sie formulierten eine Reihe von Fragen, ohne auf die Antworten des Prüflings stärker Bezug zu nehmen. In diesem Fall liefert der Prüfling Stichworte, an die der Prüfler mit einer Rückfrage oder einem neuen Thema ansetzen kann, ohne dabei die Antworten des Prüflings zu kommentieren. So zum Beispiel ein Lehrer, der dem Prüfling Fragen stellte, ihn ausreden ließ und kommentarlos weiter machte. Alles wirkte so, als würde ein Fragenkatalog abgespult, der für den Lehrer zufriedenere Antworten hervorrief. Ein Auszug aus diesem Prüfungsgespräch:²²

PR.: ... Wie kommt es denn zur Unterenwicklung?

PL.: Es gibt verschiedene Möglichkeiten, wie es dazu kommen kann, zum Beispiel, daß das Land abhängig ist von den Produkten und Preisen, zum Beispiel vom Erdöl. Und als es zur Krise kam und die Preise für Erdöl sanken, gingen auch die Einnahmen für die Länder zurück.

PR.: Welche Rolle spielen denn die Rohstoffexporte?

PL.: Es gibt einmal die Möglichkeit, daß ein Land keine Rohstoffe hat, zum Beispiel daß die Agrarstruktur sehr schlecht ist; viele Produkte können nicht angebaut werden. Und aus den Bedingungen von ganz früher war eine Entwicklung für solche Länder kaum möglich.

PR.: Ein Problem der Länder war die Schuldenbelastung. Können Sie sagen, wie es dazu kam?

PL.: Diese Länder importierten sehr viel und meistens so, daß die notwendigen Lebensmittel, Auf jeden Fall war der Import weitaus größer als der Export, ...

Kennzeichnend für diese Form ist, daß sich die prüfende Lehrperson lediglich als Stimulator von

Herbert Kalthoff: Das Zensurenpanoptikum

Antworten verhält, mit den Rückfragen aber keine Aussagen darüber macht, wie sie die Antwort des Prüflings beurteilt.

Anderer Lehrpersonen organisieren das Prüfungsgespräch in Form des – konversationsanalytisch formuliert – *cluing*-Prozesses. Hiermit wird ein Verfahren beschrieben, mit denen Lehrpersonen durch Hinweise, Stichworte oder Rückfragen den Schülern Schlüssel in die Hand geben, die die Frage (in diesem Fall die Prüfungsaufgabe) beantwortbar machen. Die Korrektur des Prüflings, d. h. die Produktion einer adäquaten Antwort, wird in Form des wiederholten Ablaufs von Lehrerfrage-Schülerantwort-Lehrerkommentar strukturiert.²³ Die Lehrpersonen, die sich dieses Verfahrens bedient haben, schienen – insbesondere bei schwächeren Schülern – möglichst noch etwas rauszuholen zu wollen. Ihre Rückfragen und Hinweise demonstrieren dem Prüfling aber auch die Inadäquatheit seiner Antworten.

Die Differenz beider Prüfungsformen macht sich daran fest, ob und in welchem Umfang dem Prüfling so fort Feedback gegeben bzw. nicht gegeben wird. Dies hängt wiederum mit der Richtigkeit der Antworten zusammen. Prüflingen, die auf die Fragen des Prüflers adäquat antworten können, und die somit im Sinne der Prüfungszeremonie prüfbar, *gut* prüfbar sind, wird eine wie auch immer chiffrierte Beurteilung ihrer Leistung vorenthalten. Das heißt auch, daß schwächere Prüflinge häufiger Rückmeldungen bekommen, die allerdings in der Notwendigkeit negativ gegen sie zu Buche schlagen, da diese als „Hilfen“ interpretiert werden. Kann die Beurteilung hinausgezögert und auf diese Weise abgedunkelt werden, so schneidet der Prüfling mit einer besseren Note ab.

Das Frage-Antwort-Spiel wird von den prüfenden Lehrpersonen als Verfahren eingesetzt, um Stärken oder Schwächen des Prüflings für die beobachtenden Lehrpersonen – das anonyme Publikum – sichtbar zu machen; der Einblick in die situative Leistung des Prüflings ist die Grundlage für die im Anschluß an die mündliche Abiturprüfung stattfindende Notenfindung. Vorher aber hat der Prüfling den Raum verlassen; die Lehrpersonen sind unter sich. Aus dem anonymen Publikum werden

nun erkennbare Personen mit distinkten Positionen.

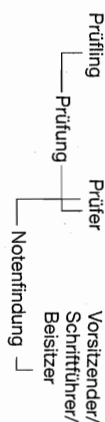


Abb. 2 Prüfungs-Paare.

Das Schema zeigt, daß die Prüfungsorganisation unterschiedliche Abstände organisiert; die Konstellation der Notenfindung kontrastiert diejenige der Prüfung. Während der Prüfler mit dem Prüfling *spricht* (Prüfungsgespräch), *beobachtet* die anderen Lehrpersonen das Geschehen; in der Person des Prüflers überschneidet sich das Geschehen. Die mündliche Abiturprüfung wird durch Identifizierungen strukturiert, die sich insbesondere auf die innere Konstellation (das Paar Prüfler-Prüfling) beziehen. So werden Prüfler und Prüfling wechselseitig identifiziert: Der prüfende Fachlehrer wird mit dem Prüfling identifiziert, wenn er Wohlwollen oder Mitleid zeigt und 'leichte' Fragen stellt; der Prüfling ist mit ihm identifiziert, da er seine Rechte übernimmt. So läßt sich formulieren, daß in der mündlichen Prüfung ein Austausch von Personen und Positionen stattfindet, und zwar die Position des Wissenden und des Lernenden. Die anderen Prüfler werden mit Lernzielen identifiziert, die sie vertreten, oder auch mit Positionen (z. B. des Direktors), die sie inne haben – Konstellationen, die als äußeres Paar bezeichnet werden können. Es ist diese spezifische Form, die die Fremd- und Selbstbeobachtung in der mündlichen Abiturprüfung annimmt.

In der Notenberatung wird zunächst von einer Lehrperson der Ablauf der Prüfung, d. h. die Leistung des Prüflings und die Fragen des Prüflers, re-stimuliert. Der Darstellung folgen Kommentare und Ergänzungen, Bestätigungen oder Korrekturen von Seiten der anderen Lehrpersonen. In ihrem Diskurs eruieren Prüfler das Gewesene; es ist fast so, als müßten sie sich kollektiv Gewißheit über das verschaffen, „was da war“ oder „was da so fehlte“. Dabei betätigen sie sich mal als Spurensucher („da war null“, „da hab ich nichts finden können“), mal als Vermesser („das war bei dem anderen tiefer“), mal als Handwerker („da müßte man ja richtig bohren“) oder auch als Feinschmecker („es war manchmal son bisschen zäh dahingekaut“). Sie stellen das Können eines Prüflings fest: „beschreiben“, „benennen“, „hinschreiben“, „ausdrücken“, „anwenden“, „sehen“ oder „erkennen“

²⁰ Im Prüfungsraum sind anwesend: der Prüfling und der Fachprüfungsausschuß. Der Ausschuß setzt sich aus vier Lehrpersonen der Schule zusammen, an der die Prüfung durchgeführt wird (so z. B. in Nordrhein-Westfalen), oder zwei prüfenden Lehrpersonen einer Schule wird eine externe Lehrperson als Vorsitzender von der aufsichtsführenden Kultusbehörde hinzubeordnet (so z. B. für Baden-Württemberg). In der Regel bestehen die Prüfungsausschüsse aus Lehrpersonen des Faches. Die Unterschiede resultieren aus länderspezifischen Regelungen der Abiturprüfung. Als zwei distinkte Formen haben sich das Zentralabitur und das dezentrale Abitur etabliert (vgl. Pape 1993; Stumpf 1993).

²¹ Diese Verkehrung wurde in den beobachteten Prüfungen durch das räumliche Arrangement symbolisiert: Alle acht Prüflinge wurden auf den Lehrertisch (das Pult) gesetzt. Nicht mehr die Lehrperson schaut von ihrem Pult auf die Kohorte der Schüler, sondern der Prüfling blickt in die Gesichter der Prüfler. Man kann diese Verkehrung, wie White (1989) es getan hat, als Initiation interpretieren. ²² Die folgenden Auszüge zum Prüfungsgespräch und zur Notenberatung wurden anhand von Notizen und Erläuterungsprotokollen rekonstruiert. Als Kürzel werden in diesem Beispiel benutzt: PR.: Prüfler, PL.: Prüfling.

²³ Die Erweiterung der Paarsequenz von Frage-Antwort in eine Sequenz von Frage-Antwort-Kommentar ist von der Konversationsanalyse am Beispiel der unterrichtlichen Kommunikation empirisch belegt worden. Auf die Frage der Lehrperson antworten Schüler, auf die Antwort der Schüler folgt ein Kommentar der Lehrperson (vgl. McHoul 1985; 1990).

Können sind für Lehrpersonen feststellbare Fähigkeiten. Sie markieren die Augenblicke, wo eine adäquate Wissensdarstellung sichtbar wurde („DA hat er gezeigt, daß er's verstanden hat“). Sie ermessens das Tun des Prüflings im Vergleich zu anderen („Der C. ist in manchen Bereichen sehr viel tiefer gegangen als die G.“), und sie überprüfen den Grad der Selbständigkeit: In welchem Ausmaß bedurfte der Prüfling, für das, was er konnte, der „Hilfe“ des Prüfers („im zweiten Teil war die nicht so gut, da brauchte die Hilfen von dir“) oder anderer Mittel (z. B. der Formelsammlung). Sie überprüfen die Sprache („in einem Deutsch, das man verstehen konnte“, „der hat doch ziemlich gestottert“), die Selbsteinschätzung des Prüflings („zum Schluß fing er an zu zittern, als er merkte, es könnte doch ne fünf werden“) und sie informieren sich durch den Prüfer über Merkmale der Schullaufbahn („hat sich so durchgemogelt“, „der hängt sich seit zwei Jahren von Formel zu Formel“). Zum vollständigen Eindruck der Prüfenden gehört auch die Information, was der Prüfling „braucht“, d. h. ob er gefährdet ist durch die Prüfung zu fallen. Dieser Frage folgen Berechnungen der Lehrpersonen anhand von Umrechnungstabellen; Lehrpersonen kalkulieren bei der Festlegung der Punkte genau die Konsequenzen für einen gefährdeten Prüfling. Auch die Befindlichkeit des Prüflings („der war schon ziemlich nervös“, „der war ganz schön träge“) sowie seine äußere Erscheinung fließen in die Zurechnung ein. Ein Lehrer:

„Schon wie er reinkommt, ob er verunsichert ist, wie er angezogen ist. Ich mein, daß darf man alles net sagen, es is aber so. Wie sich einer gibt, weil, wie er angezogen is, heißt ja nich, daß er mit Krawatte kommt, aber ob er dieses Geschehen einer Staatsprüfung würdigen kann in irgendeiner Weise durch sein Äußeres. Man is ja erstmal konserviert. Es kann schon einer reinkommen mit Turnschuhen und allem möglichen, völlig locker. Wenn der dann gut is, dann sagt man erst recht, hoopla, interessant. Aber wenn das daneben is, dann sagt man sich, naja, das auch noch. Also, so ist es nun, leider Gottes. Man sieht auch die Ausgeregtheit, also die ja wahnsinnig aufgeregt, die können manchmal kaum richtig n Satz zu Ende führen. Aber man guckt dann, naja, es war aber doch inhaltlich so dieses und so und soviel da, um dem menschlich gerecht zu werden.“

Lehrpersonen überprüfen die Symptomatik des Prüfungs und setzen sie in Relation zu dem, was er „bringt“. In diesem Vergleich wird die Würdigung der Prüfungssituation durch den Prüfling mit seiner inhaltlichen Kompetenz aufgerechnet. Wie der Prüfling die Prüfungssituation symbolisch oder

durch Kompetenz „würdigen“ kann, so „würdig“ Lehrpersonen seine „Leistung“. Das Nichtwissen von Prüflingen kann für prüfende Lehrpersonen regelrecht erwerletzend sein. So in einem Fall, in dem der Prüfling mit „sechs“ beurteilt wurde, aber aus Gründen der Gefährdung eine „fünf“ bekam, obwohl es „an sich eine Beileidung ist, dem eine fünf zu geben“ – so der Vorsitzende während der Notenberatung. Es ist ein Prozeß gegenseitiger Anerkennung, in dem Loyalitäten gefragt sind.

Neben dem Prüfling steht aber auch der Prüfer zur Debatte. Machen Lehrpersonen das Handeln des prüfenden Kollegen zum Diskussionsgegenstand, demonstrieren sie ihr intimes Wissen um die kontextuelle Gebundenheit, d. h. Konstruiertheit der Schulleistung. Sowohl Schwierigkeitsgrad der Prüfungsaufgaben als auch die Fragen des Prüfers sind als schwer („ein schwieriger Text“, „gib das mal mit eigenen Worten wieder!“), „deine Fragen wann auch nicht ohne“) oder leicht klassifizierbar („das war alles nicht besonders schwer“). Für die prüfende Lehrperson besteht die Gefahr, daß sie ihr Handeln rechtfertigen muß, sei es wegen der Prüfungsaufgaben oder sei es wegen eines schwachen Prüflings, dann, wenn sich andere Lehrpersonen fragen, wie „der nur den Leistungskurs geschafft“ hat. Dies zeigt sich insbesondere bei der Notenberatung: über sein Vorgehen meinte ein Lehrer:

„Das is ne taktisch schwerwiegende Angelegenheit. Wenn ich prüfe, dann resümiere ich die Prüfung nochmal in einigen blumigen Worten, und dann merk ich plötzlich, okay, aha, stimmt eigentlich, würde dann doch sagen, das war sicherlich im guten Bereich anzuordnen. Und dann mach ich erstmal ne Pause und beobachte den Vorsitzenden währenddessen. Der merkt ja auch schon, wo der Hase hinläuft. Und wenn er dann so irgendeine ne Geste macht, dann bin ich mutiger und sag, ich danke schon, daß man elf vielleicht sogar zwölf Punkte geben kann. Und dann steigt auch meistens ja der Protokollant ein und sagt, finde ich eigentlich auch, was denken Sie? Dann is eigentlich schon das Ding gefähten. Und wenn da dann im Zweifelsbereich bis, dann kann nix mehr passieren. Zehn, zwölf (Punkte, H.K.), okay, aber bei einigen bleibt man bei elf. Aber wenn du natürlich Zweifelsbereich sagst, und du merkst, der wird immer finsterner, dann wirts natürlich schwieriger.“

Bei der Festlegung der Note verständigen sich Lehrpersonen darüber, wo sie den Prüfling „ansiedeln“ oder „hinstecken“. Gibt die prüfende Lehrperson zunächst nur den Notenbereich („so im Dreierbereich“) an, testet sie die mitprüfenden Lehrpersonen: erst dann folgt die genaue Festlegung der Punktzahl.

- Notenfindung 1²⁴
- V: So, jetzt der Vorschlag fürs Praktika.
 - P: Dreizehn.
 - V: Dreizehn steht.
 - B: Dreizehn.
 - S: Zwölf.
 - V: Also dreizehn.
 - P: Verloren (zu S).

Die Herstellung des Konsens im Beispiel 1 findet nach einer Kontroverse, ca. 20minütigen Diskussion zwischen Prüfer, Beisitzer und Schriftführer statt: der Vorsitzende „schläft“ – wie es ein Lehrer formuliert. Der auktorisatorische Abschluß zeigt Gewinner und Verlierer: der Verlierer wird ironisiert, da er 'falsch' gesetzt hat. Ein gewisser Opportunismus ist hier am Werke, denn es geht darum, auf der Mehrheitsseite zu stehen, d. h. den 'richtigen' Einsatz zu wissen und somit das Ergebnis zu bestimmen. Zugleich handelt es sich auch um eine Selbstinszenierung des Verfahrens, d. h. der Beliebigkeit, die im Verfahren entfallt werden kann.

Im zweiten Beispiel (Notenfindung 2) beraten Lehrpersonen nach der mündlichen Prüfung von drei Schülern im Fach Physik über die Zensur; sie gehen der Reihe nach vor. Dokumentiert ist hier die Diskussion über die Leistung des zweiten Prüflings.

- Notenfindung 2
- 1 V: ... Also, wo siedeln wir ihn an?
 - 2 P: Ich würde sagen eine drei plus.
 - 3 B: Drei plus?
 - 4 P: Ja, wenn ich so denke, was da war, dann ist mir eine drei zu hart.
 - 5 P: Ja, ja, ja.
 - 6 S: Also ich habe noch einen halben Punkt mehr.
 - 7 V: Zwei minus?
 - 8 S: Ja.
 - 9 P: Wir dürfen nicht das berücksichtigen, was ich nicht gefragt habe.
 - 11 B: Das war alles nicht besonders schwer.
 - 12 P: Gut, ich bin an der Oberfläche geblieben.
 - 13 S: Du mußt sehen, was hat er von dem, was gefordert ist. Du mußt sehen, was hat er von dem, was gefordert ist. Und da war nicht viel, was er nicht gebracht hat.
 - 15 V: Auf jeden Fall war es eine Offenbarung gegenüber dem ersten.
 - 17 B: Ich würde ihm auf glatt drei setzen.
 - 19 P: Glatt drei?
 - 20 B: Das ist ein Leistungskurs.
 - 21 P: Ja sicher, aber da war auch viel Unsicherheit. Wenn man bedenkt, der hat LK Physik gemacht, da erwartet man, daß der was mitkriegt hat. Und das hat er.
 - 25 V: Sollen wir sagen eine drei?

²⁴ Zur Erklärung der Kürzel: P = Prüfer, V = Vorsitzender, B = Beisitzer, S = Schriftführer.

- 26 B: Wenn ihr damit leben könnt. Das ist ja nur meine Stimme, aber ich sag das jetzt einfach für mich.
- 28 S: Ja was braucht der denn?
- 29 P: Der hat in beiden Leistungskursen, in dem einen eine vier minus, in dem anderen eine fünf plus.
- 31 B: Dann kommt der bei einer drei auf einundzwanzig.
- 32 zweiundzwanzig Punkte.
- 33 V: Ja wollen wir ihm eine glatte drei geben?
- 34 B: Ja.
- 35 P: Augenblick mal. Ich hatte eine drei plus vorgeschlagen.
- 36 V: Dann sind das also acht Punkte, das macht mit der vier minus einundzwanzig Punkte.
- 37 S: Hat der noch andere Prüfungen?
- 40 V: Nein, der hat Abweichungen. Wenn der hier wegen Abweichung, dann muß der in den anderen Fächern auch wegen Abweichung.
- 42 P: Nein, das ist nicht wegen Abweichung, sondern wegen fehlender Punkte in den Leistungsfächern.
- 45 In dem einen hat der vier minus, in dem anderen fünf plus.
- 46 V: Einen muß er haben, den hat er jetzt. Er hat jetzt Physik befriedigend insgesamt.
- 49 P: Häh? Ja aber wir haben doch noch nicht –
- 50 V: Ja hier die mündliche, da hat er eine drei.
- 51 P: Ach so, alles klar.
- 52 V: Ja, und das sind dann hundertacht insgesamt. ...

Das zweite Beispiel demonstriert eine kontroverse Aushandlung der Zensur im Falle eines Prüflings, dessen Bestehen des Abiturs nicht gesichert ist. Der erste Teil (Z. 1-25), d. h. bis zum ersten Vermittlungsversuch des Vorsitzenden, ist ein Wechselspiel von Notenvorschlägen (Z. 2, 6, 18) und gegenseitigen Infragestellungen (Z. 3, 11, 19), denen sich Begründungen anschließen (Z. 4-5, 9-10, 13-15, 20). Deutlich wird der unterschiedliche Beobachtungsort des Prüfers und des Beisitzers. Vertritt der Beisitzer eine abstrakte Forderung („ein Leistungskurs“, Z. 20), so argumentiert der Prüfer – unterstützt vom Schriftführer – mit einer Mischung aus Lernzielorientierung (Z. 9-10, 13-15) und Sozialdimension („viel Unsicherheit“, Z. 21-24). Prüfer und Schriftführer wollen die Beurteilung an den Wissenskontext binden, der durch die Aufgaben gestellt wurde: Man könne nicht etwas bewerten, das nicht gefragt wurde. Den Einwand des Beisitzers gegen die zu leichte Prüfung entkräften sie allerdings nicht.

Der erste Versuch des Vorsitzenden, eine Verständigung auf die niedrigere Note herbeizuführen (Z. 25), schlägt fehl. Da in der Gegenüberstellung der Notenvorschläge kein Konsens erzielt werden konnte, werden weitere Informationen über das, was der Prüfling „braucht“, gesammelt. Auf die „was braucht er“-Frage des Schriftführers (Z. 28) folgt – nach der Erläuterung des Prüfers – eine

„dann kommt er auf“–Antwort des Besitzers (Z. 31–32). Hier plaziert der Vorsitzende seinen zweiten Vermittlungsvorschlag, der sich wiederum am Vorschlag des Besitzers orientiert (Z. 33). Während der Besitzer zustimmt (Z. 34), bringt der Prüfer seinen Vorschlag wieder ins Spiel (Z. 35–36) und setzt unbeirrt die Berechnung der Punktzahl fort (Z. 37–38). Der Schriftführer untermeint einen zweiten Versuch, externe Informationen mit in den Entscheidungsprozess zu integrieren und hiermit die „platte drei“ zu verhindern (Z. 39). Diese können in der Diskussion dafür den Ausschlag geben, dem Prüfling die höhere oder niedrigere Note zuzusprechen. Der Verlauf der Zeilen 40–46 zeigt, daß prüfende Lehrpersonen sich in der Situation über den Kontext ihrer Entscheidung verständigen. Das Mißverständnis des Vorsitzenden („wegen Abweichungen“, Z. 40–42) wird vom Prüfer richtig gestellt („wegen fehlender Punkte“, Z. 43–46).

Im Schlußteil (Z. 47–52) setzt der Vorsitzende seinen ersten Vermittlungsvorschlag, d. h. den Vorschlag des Besitzers, durch. Auf seine Entscheidungsalternative („einen hat er jetzt“, Z. 47–48) hin, reagiert der Prüfer mit einem erneuten Versuch, seinen eigenen Vorschlag Geltung zu verschaffen und hiermit die schlechtere Note abzuwenden. Der Satz, den er abbricht, wollte das Procedere einleiten: „Wir haben doch noch nicht (ab)gestimmt, H.K.“ (Z. 49). Der Vorsitzende insistiert auf seinen Vorschlag („da hat er eine drei“, Z. 50), worauf der Prüfer mit einem Rückzug reagiert („alles klar“, Z. 51). Abgeschlossen wird die Notefindung durch den Vorsitzenden, der feststellt, daß der Prüfling mit dem Ergebnis die Mindestanforderungen erfüllt hat („hundertacht insgesamt“, Z. 52).

In dieser Notefindung wird die Paarkonstellation des mündlichen Abtuns (s. S. 117) neu konfiguriert. Zwei Paare stehen sich gegenüber: Vorsitzender/Beisitzer vs. Prüfer/Schriftführer. Beide Paare verfahren arbeitsteilig. Ein Teilnehmer des Paares ist für die inhaltliche Argumentation zuständig, ein anderer für die Vermittlungsvorschläge oder Informationstragen. Die Arbeitsteilung wird sichtbar demonstriert durch den auffälligen Rückzug des Besitzers aus dem Teil der Diskussion, der die Entscheidung bringt; sie durchzusetzen überläßt der Beisitzer nach vorheriger Relativierung seines Urteils (Z. 26–27) dem Vorsitzenden. Deutlich wird in dieser Beratung, daß der Prüfer durch seine offenkundig werdende Orientierung am Leistungsniveau des Prüflings und durch die Integration der Sozialdimension mit dem Prüfling identifiziert wird.

Die kollegiale Verteilung der Prüflinge produziert Gewinner und Verlierer unter den Lehrpersonen; kommt dies im ersten Beispiel in spielerischer Form zum Ausdruck, so im zweiten Beispiel in einem latenten Konflikt um die Beurteilungskompetenz des Prüfers. Ein Lehrer meinte hierzu:

„Also man verteidigt jetzt etwas, also einen Schüler, oder man macht einen Schüler jetzt eher schlecht. Und man hat die eigenen Schüler vor Augen, die man vielleicht vorher geprüft hat oder die man kennt. Man hat gewisse Ansichten, Freundschaften zu einem Lehrer, es werden ja da auch Parteilagen geschlagen, gegenseitige. Da gehts dann um solche Dinge.“

Die Notefindung im mündlichen Abitur unterliegt zwei weiteren Dynamiken, die den Teilnehmern kaum kontrollierbar bzw. nicht reparabel erscheinen. Zu Beginn des Prüfungstages oder zu Beginn einer Blockprüfung zeigen Lehrpersonen mit der Note. Das heißt, Lehrpersonen sind Teil eines Prozesses, der sie den ersten Prüfling „strenger“ bewerten läßt als den dritten.²⁵ Ein Lehrer:

„Der erste wird meistens strenger beurteilt als der letzte. Aber man darfs nicht mehr ändern. Wir wann schon mal so weit, daß wir gesagt haben, nach fünf sechs Prüflingen, der erste is echt schlecht wegkommen. Immer zu schlecht. ... Am Anfang is man noch frisch und scharf, nachher werden die Klängen stumpf und müder, und dann denkt man, achja, so schlecht is es ja nun auch nich, und hoffentlich is bald Pause. Und das schleift sich ab. Das is also die FRISCHE, die fehlt, und die Schärfe.“

Vergleichbar der Anpassung an die Schülerleistung während der Klausurkorrektur, senken Lehrpersonen im Verlauf der mündlichen Abiturprüfung ihre Erwartungen. Das heißt, Lehrpersonen *starrten* optimistisch, erleben aber in bestimmter Regelmäßigkeit eine Enttäuschung, die zu einer Neuformulierung des Maßstabes führt, der kollektiv verhandelt werden muß. Hinzu kommt, daß mit zunehmender Prüfungsdauer Müdigkeitserscheinungen auftreten. Während ihrer Notenberatungen sehen Lehrpersonen die Pause von Prüfungssessenen herbei. So meinte ein Vorsitzender zu seinen Kollegen: „Dankt daran, es wartet Biensstich auf euch. Wenn wir noch langamer machen, kriegen wir nichts mehr mit.“ Der willkommene Abschluß – Kaffee, Biensstich und Pausch mit den Kollegen – wirken hier als Verfahrensbeschleuniger.

Die zeitliche Achse ihrer beruflichen Biographie spielt ebenso in der Selbstbeobachtung der Prü-

fungspraxis eine Rolle. Ein „Anfang“ ist von einem „jetzt“ unterscheidbar ein „unsicheres“ Vorgehen mit niedrig angesetzten Bewertungsvorschlägen, von einem „selbstbewußteren“, „scharfe“ Bewertungen der ersten Zeit, von „wohlwollenderen“ der späteren Jahre. Auch der umgekehrte Fall tritt auf: Einem jüngeren Lehrer, dem es sichtlich schwer fiel, der Bewertung eines Prüflings mit der Note „ausreichend“ zuzustimmen, da der Prüfling hiermit die Mindestpunktzahl verfehlt hatte, machte ein erfahrener Lehrer Mut mit der Aussicht auf einen Gewinneseffekt: „Da muß man einfach mal durch. Wenn man das das erste Mal gemacht hat, dann gehts“. Einem anderen jungen Lehrer stand die Erleichterung über den guten Ausgang „seiner“ ersten Abiturprüfung auf die Stirn geschrieben: von Schulterklöpfen seiner Kollegen begleitet verließ er den Raum.

4. Schlußbetrachtung

In dem Aufsatz ging es um die Bewertungspraxis von Lehrpersonen in den unterscheidbaren Verfahren von Einzel- und Kollektivbewertung. Die Kombination von ethnographisch generiertem Material und empirischem Konstruktivismus hat den Blick auf die Prozesse des Erzeugens schulischer Urteile gelenkt. Auf diese Weise konnte zweierlei nachgezeichnet werden: Zum einen sind Lehrpersonen als „Macher“ schulischer Urteile in den Bewertungsprozess involviert, indem sie auf sich selbst reflektieren, konstruieren sie immer auch ihre eigene Leistung. Zum anderen folgt die Bewertungspraxis – obwohl in abgeschlossenen Räumen hergestellt und somit systematisch den Augen der Öffentlichkeit entzogen – nicht einer privaten Herstellungslogik, sondern setzt auf die Darstellbarkeit ihrer Rationalität. Die Selbstkonstruktion der Bewertungspraktiker kann als ein Konstruktionsmechanismus des schulischen Differenzierungsprozesses bezeichnet werden. Dieser Befund hat zur Konsequenz, die Perspektive von der normativ angelegten Idee der Leistungsmessung der Schüler – wie sie in der Bewertungsfor-schung dominiert – in die Idee eines panoptischen Systems der Fremd- und Selbstbeobachtung zu verschieben; dies ist hier in Anlehnung an Foucault (1977) geschehen.

Bleibt die Frage, welche Konsequenzen hieraus für eine Praxistheorie resultieren. Die mikrosoziologische Perspektive geht hinter den Bourdieuschen Praxisbegriff als Ausfluß der generativen Grammatik des Habitus und der Homologie von

objektiven und mentalen Strukturen zurück und kann auf diese Weise das beobachtbare Handeln der Teilnehmer aus seiner sozialstrukturellen Überdetermination herauslösen und auf ihre lokalen Erzeugungskontexte zurückführen. Der kausaltheorische Rückgriff auf soziologische Makrodaten zur Erklärung von Resultaten, die aus Aushandlungsprozessen hervorgehen (z. B. mündliche Abiturnoten), 'blackboxiert' das soziale Geschehen, seine Struktur und seine Prozesse, in die verschiedene Materialien, Annahmen, Ansprüche etc. einfließen. Die nach einem input/output Muster soziale Differenzen ratifizierende Bewertungsmaschine Bourdieus läßt genau diese Dimensionen lokaler Praktiken unangestastet.

Mit der konstruktivistischen Präferenz für das Geschehen in der sozialen Welt wirft diese Studie einen Blick in die black box „schulische Urteilspraxis“ und macht die Konstruiertheit des Bewertungsprozesses sowie die Steuerungsmechanismen und Zurechnungsalternativen der Teilnehmer sichtbar. Sie hat gezeigt, daß mehr im Spiel ist, als die Reproduktion sozialer Klassendifferenzen. Die Bourdieusche „Bewertungsmaschine“, deren Antrieb über nur einen Übersetzungsmechanismus läuft, wird durch eine komplexe Maschinerie ersetzt, in der es um das Relevant-Machen verschiedener empirischer Relationen geht, um die Dynamik der Situation sowie um die Konstitution beobachteter Phänomene. Dies hat zwei Konsequenzen: Zum einen wird jene Mischung aus theoretischer Überdetermination und empirischer Unterkomplexität abgelöst durch einen weiter gefaßten Wirklichkeitsbegriff. Zum anderen, und hiermit verknüpft, wird für einen ebenfalls erweiterten Praxisbegriff votiert, der mehr umfaßt, als die Durchführung einer Regel (vgl. Knorr-Cetina 1994, 1996a).

Zur sozialen Wirkung schulischer Beurteilungen lassen sich über den empirischen Befund hinaus weitergehende Überlegungen anstellen. Diese sind mit einem zweiten konstruktivistischen Motiv verbunden, das auf die Produktion von Ordnung im Sinne einer Objektivierung abzielt (vgl. Knorr-Cetina 1996b). Für die Urteilspraxis heißt dies, daß Lehrpersonen nicht nur einzelne Schülerleistungen bewerten, sondern eine distinkte Verteilung der Schüler über eine Skala organisieren. Diese *schulische Ordnung* fixiert soziale Relationen, schreibt sie fest oder verändert sie, und schafft die Basis für Prozesse der Identitätsbildung auf Seiten der Schüler. Über institutionelle Prozesse der Wissensabfragung werden Personen als Individuen konstituiert, d. h. sie werden unterschied-

²⁵ Ähnlich im Eiskunstlauf: Hier wissen die Teilnehmer, daß eine hohe Startnummer bessere Punktwerte impliziert.

bar und erkennbar gemacht. Im Sinne von Geertz (1987) handelt es sich um ein institutionelles Verfahren, Personen in Relation zu anderen zu kennzeichnen. Schüler werden in eine symbolische Ordnung gebracht, die sie mit Eigennamen („Einserrkandidat“) ausstattet. Sie werden zu „Repräsentanten bestimmter distinkter Kategorien von Personen“ (Geertz 1987: 137), zu *Verkörperungen* dessen, was ihnen attestiert wurde. Die schulische Ordnung liefert eine „Vorstellungsstruktur“ (Geertz), deren Wirkung über den Verteilungsmodus hinaus anhält, d. h. sie ordnet und orientiert die Wahrnehmung nicht allein in der interaktiv geteilten Welt des Klassenraumes, sondern gerade außerhalb desselben durch Eltern, Arbeitgeber und Sozialstrukturanalysen. Das heißt, die schulische Bewertung wird in ihren sozialen Geltungsanspruch zu einer Tatsache abgedichtet, die den Erzeugungskontext ignoriert.

Schüler erzielen durch die Darstellung ihres Wissens und dessen nachträglicher Bewertung Wissen über sich selbst; Klausur für Klausur wird ihr (schulisches) Selbst- und Fremdbild zusammengefügt. Denn Wissen tritt den Schülern als eine Objektivität gegenüber, die eigenständigen Kriterien von Richtigkeit genügt. Ihr Denken ist – im Unterschied zum alltäglichen Austausch – an die Experten attestierte Richtigkeit gekoppelt. Was Schüler glauben, binden sie an die Beurteilung ihrer Antworten; die objektive Welt ist so, wie sie beurteilt worden sind. Das schulische Urteil, d. h. die Erfahrung eines andauernden Beurteilens und Zensierens, kommuniziert und erzwingt einen „Sinn von Grenzen“, der identitätskonstituierend wirkt. So sammeln Schüler nicht nur Wissen als Gewißheiten, sondern ebenso sozial konstruierte Status-Gewißheiten. Betrachtet man die schulische Institution als ein pädagogisches Laboratorium, so handelt es sich um Prozesse, die die Familien (und ihre relevanten Umwelten) durch die schulischen Objekte (d. h. die Schüler) rekonfigurieren.

Literatur

Althoff, H., 1993: Überlegungen zur Leistungsbewertung bei einer Klausur. *Der Mathematikunterricht* 39 (1): 35-38.
 Amann, K., 1990: *Natürliche Expertise und künstliche Intelligenz. Eine empirische Untersuchung naturwissenschaftlicher Laborarbeit*. Dissertation, Universität Bielefeld.
 Apple, M.W. (ed.), 1982: *Cultural and economic reproduction in education. Essays on class, ideology and the State*. London u.a.: Routledge & Kegan Paul.

Barnes, S., 1985: A study of classroom pupil evaluation: The missing link in teacher education. *Journal of Teacher Education* 36: 46-49.
 Baritzky, H., 1989: Umgang mit Zensuren in allen Fächern. Leistung und Leistungsförderung. Beobachtungen, Tests, Klassenarbeiten, Zeugnisstreifen. Frankfurt a.M.: Scriptor.
 Barton, L./Walker, S., 1983: *Race, Class and Education*. London: Croom Helm.

Baumann, J./Lohr, S./Menzel, W./Messelken, H., 1979: Korrigieren, Zensieren, Beurteilen. Seelze: Friedrich Becker, G.E., 1988: *Klassenarbeiten vorbereiten, zusammenstellen, durchführen. Aus der Sicht der handlungsorientierten Didaktik*. Pädagogik 40: 24-30.
 Benner, D./Ramesger, J., 1985: *Zwischen Zifferenzensur und pädagogischen Entwicklungsbericht: Zeugnisse ohne Noten in der Grundschule*. Zeitschrift für Pädagogik 31: 151-174.
 Bez, D., 1974: Rhythmische Schwankungen als Fehler in der Notengebung bei mündlichen Prüfungen. *Psychologie in Erziehung und Unterricht* 21: 1-14.
 Birkel, P., 1978: *Mündliche Prüfungen. Zur Objektivität und Validität der Leistungsbeurteilung*. Bochum: Kamp.
 Böhm, G./Fahn, K./Goertz, R., 1994: *Verordnung über die Bildungsgänge und die Abiturprüfung in der gymnasialen Oberstufe (APO-GOS)*. Verordnung über die Abiturprüfung für Schüler an Waldorfschulen (PO-Waldorf). Kommentar für die Schulpraxis. Essen: Wingen.

Bourdieu, P., 1974: *Zur Soziologie der symbolischen Formen*. Frankfurt a.M.: Suhrkamp.
 Bourdieu, P., 1979: *Entwurf einer Theorie der Praxis auf der ethnologischen Grundlage der kabylischen Gesellschaft*. Frankfurt a.M.: Suhrkamp.
 Bourdieu, P., 1984: *Die feinen Unterschiede. Kritik der gesellschaftlichen Urteilskraft*. Frankfurt a.M.: Suhrkamp.
 Bourdieu, P., 1989: *La noblesse d'état. Grandes écoles et esprit de corps*. Paris: Minuit.
 Bourdieu, P./Passeron, J.-C., 1971: *Die Illusion der Chancengleichheit. Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Stuttgart: Klett.
 Bourdieu, P./Passeron, J.-C., 1973: *Grundlagen einer Theorie der symbolischen Gewalt*. Frankfurt a.M.: Suhrkamp.
 Bourdieu, P./Saint Martin, M. de., 1975: *Les catégories de l'indéterminé professoral. Actes de la recherche en sciences sociales* 3: 68-93 (Deutsche Ausgabe: *Die Kategorie des professoralen Verstehens*, S. 353-393 in: P. Bourdieu: *Homo academicus*. Frankfurt a.M.: Suhrkamp 1992).

Brookhart, S.M., 1993: *Teacher's Grading Practices: Meaning and Values*. *Journal of Educational Measurement* 30: 123-142.
 Carter, H.D., 1962: *How Reliable Are Good Oral Examinations?* *California Journal of Educational Research* 13: 147-153.
 Collins, R., 1979: *The Credential Society. An Historical Sociology of Education and Stratification*. New York: Academic Press.

Cookson, P.W./Persell, H., 1985: *Preparing for Power: America's Elite Boarding Schools*. New York: Basic Books.
 Eells, W.C., 1977 [1930]: *Die Zuverlässigkeit wiederholter Benotung von aufsatzähnlichen Prüfungsarbeiten*. S. 167-176 in: K. Ingenkamp (Hrsg.) (s.u.).

Farnen, R.F., 1994: *Politik, Bildung und Paradigmenwechsel: jüngste Trends in der Kritischen Pädagogik*. In den politischen Wissenschaften, der politischen Sozialisation und in der politischen Bildung in den Vereinigten Staaten, S. 338-384 in: H. Sünker/D. Timmermann/F.-U. Kolbe (Hrsg.): *Bildung, Gesellschaft, soziale Ungleichheit. Internationale Beiträge zur Bildungssoziologie*. Frankfurt a.M.: Suhrkamp.
 Ferdinand, W., 1971: *Das Vortrags des Lehrers über die Leistungsfähigkeit bestimmter Schüler im Spiegel der Aufsatzensur*. *Schule und Psychologie* 18: 92-95.
 Finlayson, D.S., 1977 [1951]: *Die Zuverlässigkeit bei der Zensurierung von Aufsätzen*. S. 90-103 in: K. Ingenkamp (Hrsg.) (s.u.).
 Fischer, W.L., 1991: *Mathematische Kritik der Ziffernoten und ihrer Interpretation*. S. 225-249 in: J.G. von Hohenzollern/M. Liedtke (Hrsg.): *Schülerbeurteilung und Schulzeugnisse. Historische und systematische Aspekte*. Bad Heilbrunn/Obb.: Klinkhardt.
 Foucault, M., 1977: *Überwachen und Strafen. Die Geburt des Gefängnisses*. Frankfurt a.M.: Suhrkamp.
 Gaude, P., 1989: *Beobachten, Beurteilen und Beraten von Schülern. Schulpsychologische Hilfen für Lehrer*. Frankfurt a.M.: Diesterweg.
 Gaude, P./Tscherner, W.-P., 1971: *Objektivierete Leistungsmessung in der Schule. Einsatz informeller Tests im Leistungsdifferenzierten Unterricht*. Frankfurt a.M.: Diesterweg.

Geertz, C., 1987: *Person, Zeit und Umgestaltungen auf Bali*. S. 133-201 in: C. Geertz: *Dichte Beschreibung. Beiträge zum Verstehen kultureller Systeme*. Frankfurt a.M.: Suhrkamp.
 Geißler, R., 1987: *Soziale Schichtung und Lebenschancen in der Bundesrepublik Deutschland*. Stuttgart: Enke.
 Giroux, H.A., 1983: *Theories of Reproduction and Resistance in the New Sociology of Education: A Critical Analysis*. *Harvard Educational Review* 53: 257-293.
 Hartog, P./Rhodes, E.C., 1934: *An examination of examinations*. London: MacMillan.
 Fieckhausen, H., 1974: *Leistung und Chancengleichheit*. Göttingen: Hogrefe.
 Hinrichs, B./Mülstegen, E., 1994: *Schriftliche Arbeiten im Geographienunterricht*. Köln: Aulis.
 Hirschauer, S., 1993: *Die soziale Konstruktion der Transsexualität. Über die Medizin und den Geschlechtswechsel*. Frankfurt a.M.: Suhrkamp.
 Holmeyer, W., 1988: *„Allgemein anerkannte Beurteilungskriterien“ als schlechtliebe Beurteilungskriterien*. Berlin: Duncker & Humblot.
 Hopf, C., 1983: *Der Lehrer als Schüler. Überlegungen zur Praxis der Lehrerbeurteilung durch Schulfreie*. S. 203-218 in: H. Becker/H. von Henning (Hrsg.): *Zensuren. Lüge-Notwendigkeit-Alternativen*. Frankfurt a.M.: Ullstein.

Hurst, H./Nissen, R./Schulz, U., 1979: *Oberstufenarbeiten Englisch. Von der Textauswahl bis zur Bewertung*. Dortmund: Lesing.
 Ingenkamp, K. (Hrsg.), 1977: *Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsberichte*. Weinheim: Beltz (7. Auflage).
 Ingenkamp, K./Wolff, B., 1982: *Research in Oral Secondary School Leaving Examinations in Germany*. *British Journal of Educational Psychology* 52: 341-349.
 Kalthoff, H., 1994: *Die soziale Konstruktion der Erzeugenheit. Internationales und ihre praktische Pädagogik*. Dissertation, Universität Bielefeld.
 Kalthoff, H., 1995: *Die Erzeugung von Wissen. Zur Fabrikation von Antworten im Schulunterricht*. Zeitschrift für Pädagogik 41: 925-939.
 Klemisch, I./Osterhoff, K., 1993: *Klassenarbeiten und ihre Korrektur. Der Mathematikunterricht* 39 (1): 28-37.
 Knoche, W., 1969: *Jungen, Mädchen, Lehrer und Schulen im Zensurenvergleich*. Weinheim: Beltz.
 Knorr-Cetina, K., 1984: *Die Fabrikation von Erkenntnis. Zur Anthropologie der Wissenschaft*. Frankfurt a.M.: Suhrkamp.
 Knorr-Cetina, K., 1988: *Das naturwissenschaftliche Labor als Ort der „Verdichtung“ von Gesellschaft*. Zeitschrift für Soziologie 17: 85-101.
 Knorr-Cetina, K., 1989: *Spielarten des Konstruktivismus. Einige Notizen und Anmerkungen*. *Soziale Welt* 40: 86-96.
 Knorr-Cetina, K., 1994: *Primitive Classification and Postmodernity: Towards a Sociological Notion of Fiction*. *Theory, Culture and Society* 11: 1-22.
 Knorr-Cetina, K., 1996a: *Epistemic Cultures: How Scientists Make Sense (forthcoming)*.
 Knorr-Cetina, K., 1996b: *Theoretical Constructionism. On the Nesting of Knowledge Structures into Social Structures*. *Sociological Theory* (forthcoming).
 Köter, L./Grau, U., 1965: *Zur Bedingtheit der uneinheitlichen Benotung von Schülerarbeiten (Nachherzahlungen)*. Zeitschrift für experimentelle und angewandte Psychologie 12: 278-301.
 Krapppmann, L./Oswald, H., 1995: *Alltag der Schulkinder. Beobachtungen und Analysen von Interaktionen und Sozialbeziehungen*. Weinheim und München: Juventa.
 Kultusminister des Landes Nordrhein-Westfalen (Hrsg.), 1987: *Verordnung über die Bildungsgänge und die Abiturprüfung in der gymnasialen Oberstufe (Aus-Bildungs- und Prüfungsordnung gemäß §265 SchVG-APO-GOS)* mit Verwalungsvorschriften zur Verordnung über die Bildungsgänge und die Abiturprüfung in der gymnasialen Oberstufe (VVZAPO-GOS). Frechen: Ritterbach.

Kvale, S., 1993: *Examinations reexamined. Certification of students or certification of knowledge?* S. 215-240 in: S. Chaiklin/J. Lave (eds.): *Understanding practice. Perspectives on activity and context*. Cambridge: Cambridge University Press.
 Langer, A./Langer, H./Theimer, H., 1993: *Lehrer beobachten und beurteilen Schüler*. München: Oldenbourg.
 Latour, B./Woolgar, S., 1979: *Laboratory Life. The Construction of Scientific Facts*. London: Sage.

- Lewine, H.G./McGuire, C.H., 1970: The Validity and Reliability of Oral Examinations in Assessing Cognitive Skills in Medicine. *Journal of Educational Measurement* 7: 63-74.
- Litger, W., 1992: Die Fragwürdigkeit der Zensurengebung und die „Berichte zum Lernvorgang“ der Bielfelder Laborschule. *Neue Sammlung* 32: 387-404.
- Luhmann, N./Schott, K.E., 1988 [1979]: Reflexionsprobleme im Erziehungssystem. Frankfurt a.M.: Suhrkamp.
- Lynch, M., 1985: Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory. London: Routledge & Kegan Paul.
- Lynch, M., 1993: Scientific practice and ordinary action. *Ethnomethodology and social studies of science*. Cambridge: Cambridge University Press.
- Mackay, R., 1975: Conceptions of Children and Models of Socialization. S. 180-193 in: R. Turner (ed.): *Ethnomethodology. Selected Readings*. Harmondsworth: Penguin.
- McHoul, A.W., 1985: Two aspects of classroom interaction, turntaking and correction: A research report. *Australian Journal of Human Communication Disorders* 13: 53-64.
- McHoul, A.W., 1990: The organization of repair in classroom talk. *Language in Society* 19: 349-377.
- Müller, R.M., 1989: Fremdsprachenlehrer als Trainer. Oder: Was heißt „eine sprachliche Regelung geleitet haben“? Praxis des neu sprachlichen Unterrichts 36: 15-21.
- Niehus, N., 1976: *Schul- und Prüfungsrecht*. München: Beck.
- Orth, B., 1974: Einführung in die Theorie des Messens. Stuttgart: Kohlhammer.
- Pape, B. von, 1993: Die schriftliche Abiturprüfung im Fach Mathematik in Niedersachsen. *Der Mathematikunterricht* 39 (1): 67-76.
- Rosenberger, E./Lindner, K., 1984: *Handbuch der Notengebung*. Die Schülerbeurteilung in der Praxis. Zur neuen Verordnungsgebung über die Notengebung. Rangendingen: Lipura.
- Schiffner, L., 1991: *Klassenarbeiten – leichter für Lehrer, effektiver für Schüler*. Praxis des neu sprachlichen Unterrichts 38: 81-86.
- Schoiz, G., 1994: Die Konstruktion des Kindes. Über Kinder und Kindheit. Opladen: Westdeutscher Verlag.
- Schott, K.E., 1990: Erziehung als Periode. Über die Organisation von Anfang und Ende. S. 112-132 in: N. Luhmann, K.E. Schott (Hrsg.): *Zwischen Anfang und Ende*. Frankfurt a.M.: Suhrkamp.
- Schrems, E./Wolf, J., 1985: Die mündliche Abiturprüfung im Fach Gemeinschaftskunde. Hildesheim: Niedersächsisches Landesinstitut für Lehrerfortbildung.
- Schröder, H., 1974: Leistungsmessung und Schülerbeurteilung. Stuttgart: Klett.
- Schröder, H., 1990: Leistung in der Schule. Begründung, Forderung, Beurteilung. München: Arndt.
- Schröder, G., 1971: Die ungerechte Aufsatzzensur. *Bochim: Kamp*.
- Schröter, G., 1972: Informations- und Trainingsmappe für das Zensieren von Schüleraufsätzen. 617 Aufsätze von Grund- und Hauptschülern – zensiert und kommentiert. Bochum: Kamp.
- Sommer, W., 1983: Bewährung des Lehrerurteils. Eine empirische Untersuchung über den Ausgewert des Lehrerurteils für den Bildungs- und Berufserfolg. *Bad Heilbrunn/Obb.: Klinkhardt*.
- Speight, S., 1991: Would you have marked it wrong? Praxis des neu sprachlichen Unterrichts 38: 171, 176-177.
- Starch, D./Elliot, E.C., 1977 [1913]: Die Verlässlichkeit der Zensuren von Mathematikarbeiten. S. 81-89 in: K. Ingeknamp (Hrsg.) (s.o.).
- Stein, O., 1989: Wider das Korrigierenwesen. Praxis des neu sprachlichen Unterrichts 36: 265-269.
- Stemhal, H., 1983: Über Zensuren: mit und gegen Hellmut Becker nachgedacht. S. 33-52 in: H. Becker/H. von Henig (Hrsg.): *Zensuren. Lüge/Notwendigkeit/Alternativen*. Frankfurt a.M.: Ullstein.
- Stiggins, R.J./Conklin, N.F., 1992: In teachers' hands: Investigating the practices of classroom assessment. Albany: SUNY Press.
- Stumpf, B., 1993: Für und Wider zum Zentralabitur unter besonderer Berücksichtigung des Verfahrens in Baden-Württemberg. *Der Mathematikunterricht* 39 (1): 59-66.
- Streeck, J., 1979: Sandwich. Good for you. Zur pragmatischen und konversationellen Analyse von Bewerbungen im institutionellen Diskurs der Schule. S. 235-257 in: J. Dittmann (Hrsg.): *Arbeiten zur Konversationsanalyse*. Tübingen: Niemeyer.
- Terhart, E., 1979: *Ethnographische Schulforschung in den USA*. Ein Literaturbericht. *Zeitschrift für Pädagogik* 25: 291-306.
- Ushofer, R., 1949: Zur Beurteilung von Reifeprüfungsaufsätzen. *Der Deutschunterricht* 1: 84-102.
- Weiss, R., 1965: *Zensur und Zeugnis*. Beiträge zu einer Kritik der Zuverlässigkeit und Zweckmäßigkeit der Ziffernbeurteilung. Linz: Haslinger.
- White, J.J., 1989: Student Teaching as a Rite of Passage. *Anthropology and Education Quarterly* 20: 177-195.
- Wiesnerkowska, W./Nickel, H., 1974: Einflüsse auf die Beurteilung von Schüleraufsätzen. Ergebnisse einer qualitativen experimentellen Vorschulreife. S. 271-308 in: K. Heller: *Leistungsbeurteilung in der Schule*. Heidelberg: Quelle & Meyer.

Was ist soziologisch an der Wirtschaftssoziologie? Ungewißheit und die Einbettung wirtschaftlichen Handelns*

Jens Beckert

Freie Universität Berlin, John F. Kennedy Institut, Lansstr. 5-9, D-14195 Berlin

Zusammenfassung: Insbesondere in den USA hat sich die Wirtschaftssoziologie während der letzten fünfzehn Jahre zu einem herausragenden Forschungsgebiet innerhalb der Soziologie entwickelt. Der Beitrag trägt nach einem theoretischen Ausgangspunkt für die Wirtschaftssoziologie in Abgrenzung zur ökonomischen Theorie. Das Problem der Ungewißheit erweist sich als zentrale Grenze für wirtschaftswissenschaftliche Ansätze, die von der Annahme des Maximalerhandlungsansatzes von Akteuren ausgehen. Der hier vorgeschlagene Ansatz der Wirtschaftssoziologie stellt nicht die Handlungsmotive der Akteure in den Vordergrund, sondern vielmehr die Struktur der Handlungssituation. Durch Ungewißheit wird das Ordnungsproblem wieder in die Ökonomie einbezogen und es läßt sich die Frage stellen: „Wie handeln intentionale rationale Akteure, wenn sie nicht wissen, durch welche Alternativenwahl sie ihren Nutzen maximieren können?“ Mit der Verbindung von Ungewißheit mit dem Problem sozialer Ordnung kann ein systematischer Anschluß der Wirtschaftssoziologie an die soziologische Theorie hergestellt werden.

Die Wirtschaftssoziologie hat sich während der letzten fünfzehn Jahre zu einem der lebhaftesten Bereiche der Soziologie entwickelt. Insbesondere in den USA gibt es seit den achtziger Jahren einen regelrechten Boom wirtschaftssoziologischer Forschung, der seine Anfänge bereits in den siebziger Jahren mit Arbeiten unter anderem von Mark Granovetter (1973), Allan Fox (1974), Fred Hirsch (1975) und Viviana Zelizer (1979) nahm.¹ Ein wichtiges Merkmal der Forschung in diesem Bereich besteht in der thematischen, theoretischen sowie methodologischen Vielfalt der Arbeiten und dem gleichzeitigen Selbstverständnis der Autoren, dennoch zu einem gemeinsamen Forschungsziel beizutragen. Die „neue Wirtschaftssoziologie“ zielt positiv auf ein soziologisches Verständnis ökonomischer Strukturen und Prozesse und hat

ihren gemeinsamen Nenner in der Mißbilligung zentraler Elemente der Erklärung ökonomischer Phänomene in den Wirtschaftswissenschaften. Unter diesem breiten Selbstverständnis entwickeln sich verschiedene Richtungen innerhalb der neuen Wirtschaftssoziologie, die alle wirtschaftlichen Institutionen untersuchen und dabei soziale Beziehungen, Kultur, Kognition, Normen, Strukturen, Macht und Institutionen als erklärende Variablen für das Verständnis wirtschaftlicher Phänomene verwenden. Das breite, sowohl positive als auch kritische Selbstverständnis der Wirtschaftssoziologie läßt sich in dem wegweisenden Aufsatz *Economic Action and Social Structure: The Problem of Embeddedness* von Mark Granovetter (1985) erkennen. Im ersten Teil des Aufsatzes entwickelt Granovetter das Konzept der Einbettung, während sich der zweite Teil der Kritik an dem Transaktionskostensatz von Oliver Williamson widmet.

Die Heterogenität der unterschiedlichen Ansätze der Wirtschaftssoziologie läßt die Frage aufkommen, ob es möglich ist, einen theoretischen Kern für die Wirtschaftssoziologie zu identifizieren, der eine Alternative zu wirtschaftswissenschaftlichen Ansätzen bieten kann. Worin besteht der spezifische soziologische Beitrag für das Verständnis ökonomischer Phänomene? Die wohl am weitesten verbreitete Antwort auf diese Frage besteht in der soziologischen Zurückweisung der im Zentrum der ökonomischen Theorie stehenden Maximalerhandlungsannahme. Seit der Formulierung der Disziplin im letzten Jahrhundert haben sich Soziologen gegen das handlungstheoretische Modell eines in-

* Ich möchte mich bei den Teilnehmern des Workshop on economic sociology (JOIE) am Department of Sociology der Princeton University und bei den Mitgliedern des Kolloquiums Moderne soziologische Theorie von Hans Joas am John F. Kennedy Institut der FU Berlin für ihre hilfreichen Kommentare bedanken. Außerdem möchte ich Paul DiMaggio, Mark Granovetter, Hans Joas und Arthur Srinivasan für ihre ausführlichen Hinweise zu einer früheren Version des Beitrags danken. Die finanzielle Unterstützung der Forschung für diesen Aufsatz durch ein Stipendium der Gottlieb Daimler- und Karl Benz Stiftung wird dankbar vermerkt. Der vorliegende Text ist die veränderte und leicht gekürzte Fassung eines Aufsatzes, der am Ende des Jahres in *Theory and Society* erschienen wird.

¹ Für die wirtschaftssoziologische Debatte in den achtziger Jahren sei insbesondere auf die Sammelbände Granovetter/Swedberg (1992), Swedberg (1993) und Zukin/DiMaggio (1990) hingewiesen.