Zeitschrift: Bildungsforschung und Bildungspraxis : schweizerische Zeitschrift für Erziehungswissenschaft = Éducation et recherche : revue suisse des sciences de l'éducation = Educazione e ricerca : rivista svizzera di scienze dell'educazione

Herausgeber: Schweizerische Gesellschaft für Bildungsforschung

Band: 10 (1988)

Heft: 3

Artikel: Verbalzeugnisse in Primarklassen der Bundesrepublik Deutschland : Zielsetzung, Einführung, Bewährungskontrolle

Autor: Ingenkamp, Karlheinz

DOI: https://doi.org/10.5169/seals-786326

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 11.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

VERBALZEUGNISSE IN PRIMARKLASSEN DER BUNDESREPUBLIK DEUTSCHLAND
- Zielsetzung, Einführung, Bewährungskontrolle -

Karlheinz Ingenkamp

Die Länder der Bundesrepublik Deutschland haben von 1970-1977 in 1. und z.T. auch in 2. Klassen der Grundschulen Zeugnisberichte anstelle von Ziffernzensuren eingeführt. Für die Formulierung der Berichte wurden Hilfen in "Mustertexten", "Kriterienlisten" sowie Lehrgängen geboten. Die diagnostiche Kompetenz der Lehrer, bes. in der Schülerbeobachtung, wurde nicht geschult. Die mit der Einführung der Zensurenberichte verbundenen Erwartungen auf angstfreien, motivierenden Unterricht wurden von einigen Autoren bestritten. Untersuchungen zur Bewährung der Zeugnisberichte liegen kaum vor. Die wenigen Studien, so anfechtbar sie methodisch sind, bestätigen eher die Annahme, dass die administrativ verordnete Reform weder in der Schülerbeurteilung noch in der Verbesserung des Unterrichts ihre Ziele erreicht hat.

Zielsetzungen bei der Einführung des Zeugnisberichtes

Im Zusammenhang mit der angestrebten Reform der Grundschule hatte die Kultusministerkonferenz der Länder der BRD 1970 beschlossen : "In der 1. und 2. Klasse ist eine allgemeine Aussage über die Leistungen eines Kindes im Hinblick auf das Ziel dieser Schulstufe bedeutsamer als die vorgeblich genaue Benotung der Leistungen in den einzelnen Teilgebieten des Unterrichts. In diesen beiden Klassen ist daher jeweils am Ende eines Schuljahres eine allgemeine Beurteilung des Kindes in freier Form im Zeugnis zu erteilen. Neben der Begutachtung des Sozial- und Arbeitsverhaltens sind Hinweise auf Interessen, besondere Fähigkeiten und Schwächen zu geben. Dabei muss zusammenfassend festgestellt werden, ob und inwieweit die Leistungen mit der Einschätzung des geistigen Leistungsvermögens übereinstimmen...".

In den Ländern der Bundesrepublik wurden von 1970-1977 Regelungen in Kraft gesetzt, die im 1. und teilweise auch im 2. Schuljahr Zensuren und Zeugnisse abschafften und Zeugnisberichte vorschrieben, die sich häufig nicht nur auf Leistung und Arbeitsverhalten beschränkten, sondern auch Sozialverhalten einbezogen. Der Arbeitskreis Grundschule hat 1978 Empfehlungen zur weiteren Entwicklung der neuen Beurteilungsbestimmungen für die Grundschule beschlossen. Darin stellt er fest, dass die pädagogischen Intentionen der Diagnose und Förderung nur eingelöst werden können, wenn "eine differenzierte Dokumentation der individuellen Lernvoraussetzungen und Lernfortschritte" vorliegt. "Dazu bedarf der Lehrer praktikabler Instrumentarien, mit denen er die erforderlichen Daten festhalten und finden kann. Die derzeitigen Hilfen reichen bei weitem nicht aus". Im einzelnen wird dann eine verstärkte Einbeziehung der Pädagogischen Diagnostik in die Lehrerausbildung, die Entwicklung von Beobachtungsverfahren, die "Kenntnis und Anwendung sowie kritische Einschätzung von subjektiven und objektiven Mess- und Beurteilungsverfahren" und die Förderung durch Forschungsvorhaben gefordert (Arbeitskreis, 1978). Diese Forderungen sind nicht realisiert worden. Die Ziele, die mit den Verbalberichten in der Eingangstufe verbunden werden, lassen sich zu folgenden Zielgruppen zusammenfassen (vgl. z.B. Burk, 1979, Bartnitzky und Christiani, 1979, Schmack, 1977) :

A. Ermutigende Erziehung statt Leistungsdruck
B. Förderung der sozialen Kooperation statt Konkurrenzverhalten
C. Erhöhung der Chancengleichheit statt Leistungsabfall der Benachteiligten
D. Individuelle Förderung statt frontal gesteuertem Lerngleichschritt

Die Gestaltung des Zeugnisberichtes

In fast allen Bundesländern wurden ab 1975 Zeugnisberichte für die 1., oft auch für die 2. Grundschulklasse eingeführt. Am intensivsten war die Vorbereitung der Lehrer auf die neue Zeugnisform in Nordrhein-Westfalen (vgl. Christiani, 1983). Neben Fortbildungstagungen, Konferenzen u.ä. gab es schriftliche Handreichungen, die sich prinzipiell in "Mustertexte" (z.B. in Baden-Württemberg und Bayern) oder "Kriterienlisten" (z.B. in Hamburg, Niedersachsen und anfänglich auch NRW) unterteilen lassen (vgl. z.B. Bolscho, 1979a).

Aber bevor die Lehrer ihre Verbalbeurteilungen niederschreiben konnten, mussten sie ja erst einmal beobachten. Und in die Phase der Vermittlung diagnostischer Kompetenz hat man nicht genug Zeit und Mühe investiert. Bei

Christiani (1983, S. 202) heisst es dann lapidar : "Gleichzeitig stellte sich heraus, dass die Anforderungen an die pädagogisch-diagnostische Praxis (Beobachtungsbögen, Schätzskalen, informelle Tests usw.) überzogen waren... 'Der Lehrer muss beobachtetes Verhalten als Erinnerungsstütze aufschreiben. Hierfür hat sich... das bekannte und bereits bewährte pädagogische Tagebuch durchgesetzt' ". Hier liegt die Schwäche des Reformansatzes. Man ist nicht den schweren Weg gegangen, die diagnostische Praxis zu ändern und zu verbessern. Dafür hätte man diagnostische Instrumente praxisnah entwickeln und die Lehrer von ihren Vorzügen überzeugen müssen. Stattdessen hat man im wesentlichen nur die Mitteilungsform geändert und sich bei den wenigen Bewährungskontrollen auch vorwiegend damit begnügt, ob - wie bei Aufsätzen - die Formulierungen die richtige pädagogische Einstellung erkennen liessen. Ob die Formulierungen die Fakten gültig erfassen, interessierte kaum, wie wir sehen werden. Die Gefahr verbalen Leerlaufs war damit von Anfang an gegeben.

Das Problem liegt darin, dass der Lehrer in seiner Ausbildung nicht in Verhaltensbeobachtung geschult wurde und wenig über Beobachtungsfehler weiss. Jede Beobachtung und ihre Protokollierung erfordern mehr Aufwand als die Zuteilung von Ziffernzensuren. Es ist sehr schwer, Lehrer vom Wert dieser Zusatzarbeiten zu überzeugen, wie die Versuche mit "Entwicklungsberichten" in Rheinland-Pfalz, mit den "Hilfen für die Beobachtung von Schülern und für die Abfassung von Schülerberichten" (vgl. Bolscho, 1979b) und die Diagnosebogen in Gesamtschulen bewiesen haben. Für diese praxisnahen Arbeiten steht auch zu wenig wissenschaftliche Kapazität zur Verfügung, so dass man sich mit recht beliebigen und unerprobten Anleitungen zufrieden gibt, wie z.B. dem "Beobachtungsbogen" von Firnkes (1984, S. 241).

In den Mustertexten wurden dann wünschenswerte und nicht wünschenswerte Beispiele gegenübergestellt, wie z.B. hier im Amtsblatt 17/1977 des Kultusministeriums Baden-Württemberg für den Bereich "Lernen" :

Wünschenswert :

Jürgen hat den Leselehrgang erfolgreich abgeschlossen. Fremde einfache Texte las er fliessend. Er konnte auch den Sinn des Gelesenen erfassen. Einfache Sätze konnte er aus dem Gedächtnis richtig aufschreiben. Er konnte Mengen darstellen und beschreiben sowie im Zahlenraum bis 20 sicher zu- und abzählen. Er beherrschte einfache Regeln des Strassenverkehrs und hat sich Grundkenntnisse über seine Umwelt angeeignet. Besonderes Interesse zeigte er an Bilderbüchern mit technischem Inhalt.

+ = häufig
/ = hin und wieder
- = selten

		Zeitpunkt	Bes. Situationen
Kontakt-fähigkeit	geht auf Mitschüler ein wirkt ausgleichend kann um Hilfe bitten hat Freunde		kann zuhören
Kooperation	arbeitet mit in PA/GA arbeitet an gemeinsamen Planungen mit kann sich absprechen stellt eigene Interessen zurück		
Kritik-fähigkeit	Vertritt eigene Auffassung Erkennt Argumente an begründet eigenes, nicht angepasstes Verhalten		
Leistungs-bereitschaft	lernt ohne ständige Anstösse arbeitet regelmässig mit ist für bestimmte Sachen besonders zu interessieren erledigt Aufgaben zügig		
Selbständig-keit/Konzen-tration	kann Aufträge aufnehmen und umsetzen kann planvoll vorgehen ist nicht ständig auf Hilfen angewiesen bringt Arbeiten ohne Drängen zu Ende kann sich über längere Zeiträume hinweg konzentrieren		
Zuverlässig-keit	vergisst Arbeitsmaterial nicht arbeitet gewissenhaft verantwortungsbewusst bei eigenen/fremden Sachen hält sich an Vereinbarungen hält ihm übertragene Aufgaben durch		
Denk- und Vorstellungs-vermögen (Kreativität)	stellt Fragen hat Ideen findet Lösungswege verarbeitet Gelerntes macht Verbesserungsvor-schläge findet sich schnell in neuen Aufgabenstellungen zurecht erkennt Zusammenhänge		

Nicht wünschenswert :

Im Lesen ist Jürgen sehr gut und den anderen weit voraus. In Mathematik ist er gut. Sein interessiertes Mitarbeiten im Sachunterricht ist lobenswert."

Der Vergleich zeigt zunächst, dass die wünschenswerte Fassung umfangreicher ist und auf Vergleiche zu anderen Schülern verzichtet. Sie gibt mehr Informationen über Lerninhalte, aber sie ist natürlich auch nicht eindeutig. Was "einfache" Sätze oder Regeln sind, was "sicher" und "Grundkenntnisse" bedeuten, ist sehr interpretationsbedürftig.

In einer Kriterienliste wird der Bereich "Sprache" für das 1.Schuljahr z.B. folgendermassen abgedeckt :

Leselehrgang

- kennt (nur wenige, die meisten, alle) Buchstaben ;
- kann (geübte, bekannte, neue) Wörter lesen ;
- kann (geübte, ungeübte) Texte in Druck- und Schreibschrift lesen ;

Schreiben

- schreibt (unleserlich, lesbar, klar, gegliedert) ;
- schreibt (verkrampft, hastig, langsam, flüssig) ;
- verwechselt noch Buchstaben, lässt noch Buchstaben aus ;
- kann nicht/kann geübte Buchstaben/Wörter/kurze Sätze nach Vorlage/aus der Vorstellung schreiben ;

Sprachfähigkeit

- kann selbst einfache Vorgänge/Zusammenhänge sprachlich noch nicht ausdrücken ; (damit ist das Ziel der Klasse hinsichtlich der Sprachfähigkeit noch nicht erreicht).
- spricht (undeutlich, stockend, deutlich, fliessend) ;
- verfügt über einen (begrenzten, einfachen, umfangreichen) Wortschatz ;
- spricht in (unvollständigen, einfachen, zusammenhängenden) Sätzen ;
- spricht (...) fehlerfrei ;
- erzählt/berichtet/beschreibt (ungenau, umständlich, einfach, treffend, gewandt).

Auch hier bleiben viele der angebotenen Begriffe sehr interpretationsbedürftig, z.B. "wenige", "geübt", "hastig", "nicht ausdrücken". Wer Klasse und Unterricht nicht gut kennt, kann aus einem nach derartigen Formulierungshilfen angefertigten Verbalbericht kaum ein zutreffendes Bild gewinnen.

Die Zielsetzung der Zeugnisberichte vor dem Hintergrund der Forschung

Die Literatur zur Einführung der Zeugnisberichte liest sich wie Erweckungsliteratur. Sie sieht das pädagogische Paradies ganz nah, wenn man nur den Teufel "Zensur" vertreibt. Stützt der Forschungstand solchen gläubigen Optimismus ?

Der "ermutigenden Erziehung" wird in allen Zieldiskussionen hoher Stellenwert eingeräumt. Die "zensurenfreie Beurteilung" wird als "ein entscheidendes Vehikel" zur Annäherung an diese Konzeption gesehen, denn "die zensurenbezogene Leistungsbewertung steht quer zum Förderauftrag der Grundschule" (Christiani, 1983, S. 197 ff.). Dagegen ist die Zensur als sozialbezogene Beurteilungsnorm "ein Übel", das "den pädagogischen Leitideen der Grundschule entgegenwirkt" und zwangsläufig Versager produziert, "weil es statistisch bedingt immer Schüler gibt, die von der Durchschnittsleistung nach unten abweichen" (Müssener, 1984, S. 136). Für diese ermutigenden Mitteilungen soll nur die individuelle Bezugsnorm, auch temporale genannt, geeignet sein. Später soll dann auch die kriterienorientierte Bezugsnorm hinzutreten. Aber die "sozialbezogene" Beurteilungsnorm müsste "längst aus allen Schulstufen und -formen verbannt sein" (Müssener, 1984, S. 136).

Der "niederdrückende Vergleich mit anderen" (Firnkes, 1984, S. 241) wird immer betont, aber in dieser Literatur wird nie eine Zensurenstatistik der ersten Schuljahre angeführt, die angibt, wieviel Schulanfänger durch schlechte Zensuren belastet werden könnten. Seyfried (1983, S. 10) teilt am Beispiel grosser Stichproben mit, dass in der Gesamtnote am Ende der 1.Klasse 19,5 - 30,3 % eine 1, aber nur 2,9 - 3,9 % eine 5 auf der fünfstufigen österreichischen Notenskala erhalten. Auch die 2 hat mit 36,4 % einen sehr viel höheren Anteil als die 4 mit 9,3 - 11,7 %.

Man kann nun sagen, dass auch 3 - 4 % mit sehr schlechten und bis zu 12 % mit schwachen Zensuren zu viel Entmutigte seien (sofern sie sich entmutigen lassen und nicht Ausweichmechanismen einsetzen). Aber die Änderung der

Mitteilungsform hebt die sozialen Vergleichsprozesse und damit die Entmutigung nicht auf. Auch wenn man die Bildchen, Sternchen u.ä., die vielen Lehrern als Zensurenersatz dienen, als Irrweg verurteilt, kann man nicht übersehen, dass die Kinder in dieser Leistungsgesellschaft ihre Lehrer ständig auf differenzierende Anerkennung hin beobachten. Glaubt man wirklich, man könne die berufserfahrenen Eltern davon abbringen, die Zensurenberichte wie Personalbeurteilungen auf verborgenen Tadel hin abzuklopfen ? Solche Fragen stellen, heisst nicht, dem Leistungskampf aller gegen alle in der Eingangsstufe das Wort zu reden. Aber es heisst, eine Änderung der Mitteilungsform noch nicht als Pädagogik anzusehen.

Immer wieder taucht bei den Befürwortern der Verbalzeugnisse die Erwartung auf, die intraindividuelle (temporale) und später die kriterienorientierte (sachliche) Bezugsnorm würden Konkurrenz- und Leistungsdruck mindern, sachorientiert motivieren und ermutigend wirken. Nirgendwo wird nachgewiesen, dass die Befürworter das durch eigene oder fremde Untersuchungen zu belegen versucht hätten. Der Forschungsstand ist ja keineswegs so, dass sich die generalisierenden Behauptungen daraus ableiten liessen. Lissmann fasst ihn (1984, S. 12) folgendermassen kurz zusammen.

"1. Noten mit sozialer Bezugsnorm haben vor allem auf leistungsfähigere Lernende günstige Auswirkungen.
2. Bei normorientierten Tests fallen die Leistungen höher aus als bei lehrzielorientierten Tests.
3. Lehrzielorientierte Tests erzeugen dagegen eine positive Einstellung zum Lehrstoff".

Wenn man die Wechselwirkungen von Norm und Häufigkeit der Rückmeldung z.B. hinzunimmt, wird das Bild verwirrender. Nie bestätigt es den Optimismus der Zeugnisreformer.

In der zweiten Zielgruppe steht die Forderung nach Abbau des rivalisierenden Lernens und die Förderung des Miteinanderlernens im Mittelpunkt. Das ist ein erstrebenswertes Ziel ! Aber um als Lehrer in den komplizierten Prozess sozialer Interaktionen in der Schulklasse erfolgreich eingreifen zu können, bedarf es sehr intensiver Kombinationen von diagnostischen und therapeutischen Bemühungen, wie erst wieder Petillon, 1982, in seinen umfassenden theoretischen und empirischen Analysen gezeigt hat. Wenn wir annehmen, wir könnten hier über eine Änderung der Mitteilungsform schon etwas erreichen, dann lenkt uns das eher von den eigentlichen Aufgaben ab.

Bartnitzky und Christiani (1979, S. 16) stellten ferner fest : "Dem Auftrag der Grundschule, ein Mehr an Chancengleichheit herzustellen, widerspricht es, Leistungsunterschiede nach der sozialen Bezugsnorm festzustellen". Die Schule "hat die Aufgabe, den ungleichen Lernvoraussetzungen mit ungleichen Lernangeboten und -anforderungen zu begegnen." Auch zur Erreichung dieses Ziels soll der Ersatz von Zensuren durch Verbalberichte erheblich beitragen.

Schmidt und Silkenbeumer (1983) halten das schlicht für "Ideologie" (S. 215) und behaupten, dass durch den Beobachtungsbogen, der zur Vorbereitung des Verbalberichtes geführt wird, "ein Prozess in Gang gesetzt wird, der allmählich eine Eigendynamik entwickelt, die den schichtenspezifischen Selektionsprozess verstärkt und die 'deviante Karriere' - vor allem für den Schüler aus der Unterschicht - fast unausweichlich werden lässt" (S. 213). Auch dies ist eine völlig unbewiesene Behauptung. Aber diese Autoren sehen wenigstens, dass sich in dem Prozess von der Wahrnehmung bis zur Niederschrift im Zensurenbericht implizite Persönlichkeitstheorien der Lehrer auswirken können. Bartnitzky und Christiani (1979, S. 76) geben zwar zu : "Der Schülerbeobachtungsbogen garantiert keinen Schutz vor allen möglichen Beobachtungsfehlern und Beurteilungsfehlern." In dem für diese Literatur so typischen Wunschdenken fahren sie fort : "Deshalb muss sich der Lehrer selbst ständig prüfen, ob er nicht beim Ausfüllen des Bogens eigene Wertvorstellungen, Verhaltenserwartungen, Gefühle oder Vorurteile einfliessen lässt." Als ob das helfen könnte !

Ch. Schwarzer (1977, S. 65) hat in einer Untersuchung an 361 Erstklässlern errechnen können, dass 17,3 % der Streuung des Lehrerurteils durch die Sozialschichtzugehörigkeit der Kinder aufgeklärt werden. Diese soziale Abhängigkeit des Lehrerurteils wird nicht gemildert oder gar aufgehoben, wenn man statt Ziffern verbale Beschreibungen fordert. 11-16 Jahre nach Einführung der ziffernlosen Zeugnisse dürfte man erwarten, dass Untersuchungen vorgelegt werden, die nachweisen, dass die mit der Umstellung vom Ziffernzeugnis zum Verbalbericht ausgesprochenen Erwartungen realisiert werden können.

Untersuchungen zur Bewährung der Zensurenberichte

Aus Nordrhein-Westfalen (NRW) sind vier Untersuchungen bekannt, die in der Reihenfolge der Publikation (Schmack, 1978 ; R. Schwarzer, 1980 ; Benner/Ramseger, 1985 ; Scheerer u.a. 1985) angeführt werden sollen.

Schmack liefert eine Untersuchungsskizze im Anfangsstadium der neuen Zeugnispraxis, als noch Formulierungshilfen des Ministeriums NRW angeboten wurden. Er hat 110 Zeugnisse von 14 Bochumer Lehrkräften vorwiegend impressionistisch ausgewertet. Dabei sollen Zeugnisniederschriften überwogen haben, die in möglichst knapper Form in Anlehnung an die Formulierungshilfen Aussagen zu den drei geforderten Bereichen Sozialverhalten, Arbeitsverhalten, Hinweise zu Lernbereichen bringen. Nur "in einer geringeren Zahl fällt eine intensivere, breitere und differenziertere Fassung" (1978, S. 241) auf. Die Zeugnisse hatten Umfänge von 60-190 Wörtern, "im Mittel dürfte es sich um ca. 100 Wörter handeln" (S. 241). Auch bei kurzen Zeugnissen konnten alle Aspekte berücksichtigt sein, es konnte aber auch sein, dass zu wenig positive Aussagen zu erwarten waren. Zum Sozialverhalten fanden sich "meist 3 bis 4 Aussagen", beim Arbeitsverhalten war "die Aussagendichte nachweislich höher", und der Lernbereich war "am breitesten und differenziertesten nach den Lernständen beschrieben" (S. 242). Aber nur in 13 von den 110 Zeugnissen waren Leistungen ausserhalb von Sprache und Mathematik erwähnt worden. Schmack bringt dann illustrierend Beispiele aus Verbalzeugnissen. Er hat ausserdem Protokollnotizen der Lehrer über das Gespräch mit den Eltern bei der Zeugnisübergabe herangezogen, ohne allerdings Effekte der sozialen Erwünschtheit in den Aussagen der Eltern und der Verzerrung durch Bestätigungserwartung der Lehrer zu diskutieren. 60,9 % der Eltern haben demnach Zustimmung und Zufriedenheit geäussert. Nur 0,9 % votierten eindeutig negativ (S. 247). Erfahrungen aus weiteren 97 Zeugnissen sollen das Bild bestätigen, ohne dass Details mitgeteilt werden. Schmack (1978, S. 253) will auf Grund dieser sehr fragmentarischen Studie noch nicht prognostizieren, "inwieweit die neue Zeugnisform 'Leistungsdruck' auf die Dauer abbauen wird, ob eine solche Massnahme dazu beitragen kann, dieses Phänomen zu relativieren und den Schulanfang zu einem 'stressfreien' Raum zu gestalten."

Methodisch anspruchsvoller ist die Querschnittsuntersuchung von R. Schwarzer (1980). Er hatte eine Stichprobe von 135 Schülern(-innen) aus 6 Klassen in drei Schulen. Die Kinder waren gerade in die 2. Klasse übergewechselt. Die Untersuchung verfolgte zwei Fragestellungen :

1. Wird in den Zeugnisberichten die soziale Bezugsnorm vermieden und verbale oder adverbale Sprache gegenüber substantivischer bevorzugt ?
2. Welche Einstellung haben die Eltern zum neuen Zeugnis ?

Die Zeugnisse wurden einer quantitativen Inhaltsanalyse unterzogen, wobei für die beiden Aussageeinheiten "Lernbereiche" und "Arbeits- und Sozialverhalten" ein Schema mit 11 Kategorien zur Einschätzung entwickelt wurde.

Jede Aussageeinheit in jedem Zeugnis musste nach den 11 Kategorien eingestuft werden, die von "1. Ich-Botschaften über 4. Temporaler Vergleich, 5. Kriterialer Vergleich... 8. Verbale/adverbale Information bis zu 11. Werthaltige Negativaussagen" reichten. Für jede Analysekategorie wurden Beispiele gegeben und vier Ausprägungsgrade vorgeschrieben. Über Kodiererübereinstimmung wird nichts gesagt.

Der Elternfragebogen enthielt 13 Fragen, wie z.B. : "War das Zeugnis für Ihr Kind verständlich genug ?", "Haben Sie beim Lesen des Zeugnisses versucht, die Formulierungen in Noten umzuwandeln ?", die jeweils beiden Eltern vorgelegt wurden. Mit den 6 Lehrerinnen wurden offene Interviews durchgeführt. Der Unterricht wurde einmal beobachtet, wobei nur darauf geachtet wurde, welche Bezugsnorm die Lehrer wählten und ob und welche Eigenschaften sie möglicherweise Schülern zuschrieben.

Die Zeugnisse enthielten 49 bis 161, durchschnittlich 100 Wörter. Die Unterschiede sind auf Lehrereffekte zurückzuführen. Zwischen Zeugnisumfang und sozialen Vergleichen bestehen keine, mit temporalen und kriterialen Bezugsnormen schwach positive Zusammenhänge, "die dafür sprechen, dass mit erhöhter Textlänge auch mehr ermutigende Äusserungen vorgenommen werden" (S. 10). Da entmutigende Äusserungen nicht zugelassen waren, ist das durchaus erwartungsgemäss. Die Eltern reagieren auf die Wortlänge nicht unterschiedlich. Keines der 135 Zeugnisse war in Briefform mit "Du-Anreden" geschrieben. Soziale Vergleiche, wie z.B. "überdurchschnittlich", tauchten im Lernbereich in zwei Zeugnissen, beim Arbeits- und Sozialverhalten in einem Zeugnis auf. Temporale Vergleiche, also die individuelle Bezugsnorm, fanden sich beim Arbeits- und Sozialverhalten in 11, bei den Lernbereichen in 32 Fällen. Kriteriale Vergleiche waren im Sozialbereich sehr häufig, bei mindestens 116 Zeugnissen, im Lernbereich sehr selten, bei mindestens 3 Zeugnissen. Die letzten Aussagen sind so vage, weil Schwarzer leider zwischen den Bezugsnormen die Häufigkeitseinteilung wechselte und damit den exakten Vergleich unmöglich machte. Bei der seltenen Anwendung individueller und sachlicher Bezugsnorm im Lernbereich ist Schwarzers (1980, S. 12) Bewertung anzuzweifeln, der darin eine "konzepttreue Verwirklichung der Zeugnisneuregelung" sieht. Wir wissen inzwischen, wenn auch aus anderen Altersstufen, dass Lehrer bei freien mündlichen Rückmeldungen ganz selten die soziale Bezugsnorm wählen. Selbst wenn sie ausdrücklich dazu aufgefordert werden, tun sie das im Mittel nur 1,05 mal pro Stunde (Paetzold, 1982, S. 254 f.).

Die substantivischen Urteile waren beim Arbeits- und Sozialverhalten sehr gering, aber "je mehr jemand zu einer verbalen Beschreibungsweise des

Arbeits- und Sozialverhaltens tendiert, desto weniger tendiert er zu einer entsprechenden Darstellung der Lernbereiche" (Schwarzer, 1980, S. 17). Das und andere Inkonsistenzen deuten eigentlich darauf hin, dass die geforderte Beurteilung der Gesamtpersönlichkeit noch nicht gelungen ist. Positive und negative Wertungen traten in den Zeugnissen etwa gleich selten auf. Den Richtlinien nach hätte man ein deutliches Überwiegen positiver Wertungen erwartet.

Die Zustimmung der Eltern bewegt sich auch zu dieser Zeugnisform in dem bekannt hohen Rahmen. Schwarzer sieht darin den Einfluss sozialer Erwünschtheit. Geringer als üblich ist die Zustimmung zu den Fragen, ob das Zeugnis für das Kind verständlich genug war und ob das Kind durch das Zeugnis ermutigt wurde. Die Eltern haben zu 45 bzw. 36 % versucht, die Formulierungen in Noten zu übersetzen. Nur 32 bzw. 34 % waren der Meinung, man solle den zensurenfreien Raum auch auf die 3.Klasse ausdehnen.

Es bestehen positive Korrelationen zwischen dem Leistungsniveau der Kinder und der Zusammenarbeit Eltern-Lehrerin (0.48), zwischen der Zusammenarbeit und der Wahrnehmung einer Ermutigung des Kindes (0.25-0.27) sowie der Lernschwäche des Kindes und der Schwierigkeit, die die Zeugnisformulierung der Lehrerin bereitet (0.61) (Schwarzer, 1980, S. 24). Die Lehrerinnen fühlten sich aber ungeachtet der Schwierigkeiten in der Lage, individuelle Zeugnisse für jeden Schüler zu erstellen. Schwarzer beurteilt die untersuchte Praxis günstig, hält aber weitere Untersuchungen mit aufwendigeren Methoden für notwendig. Es sollte auch nicht übersehen werden, dass der Untersuchungsansatz keine Aussagen zur Validität der Zeugnisse erlaubte.

Ebenfalls eine quantitative Inhaltsanalyse, in diesem Fall eine computerunterstützte Analyse mit dem Programmpaket TEXPACK, haben kürzlich Scheerer u.a. (1985) vorgelegt. Sie beschränken sich allerdings auf den Bereich "Arbeits- und Sozialverhalten". Die Stichprobenauswahl von 1 840 Zeugnissen erfolgte stufenweise, wobei offenbleibt, ob dabei randomisiert wurde. Für 32 beteiligte Klassen war ein Längsschnittvergleich von der 1. bis zur 3. Klasse möglich. Für 63 bzw. 93 Klassen sind Quervergleiche auf der 1. bzw. 2. Klassenstufe durchführbar. Die Repräsentativität der Stichprobe soll über Merkmale der beteiligten Lehrerinnen nachgewiesen werden, was nicht überzeugt. Für das bei TESTPACK erforderliche Kategoriensystem griffen die Autoren auf die Formulierungshilfen des Kultusministeriums von 1976 und den Raster von Bartnitzky/Christiani (1977 u. 1979) zurück.

Bei der Auszählung der Sätze in den Zeugnissen war eine signifikante Abnahme von der 1. zur 2. Klasse und ebenfalls für die Klassen festzustellen, die im 2.Jahr nach der Reform untersucht wurden.

Es ist unmöglich und auch nicht nützlich, hier auf die Auszählungen der Inhaltskategorien einzugehen. Hier bleibt nur Raum für zusammenfassendere Aussagen. Die Autoren belegen, dass nur etwa 4 % der Zeugnisse "ausdrücklich normative Sätze in Form von Appellen oder Aussagen zur Entwicklung des Schülers" enthalten, während "die überwiegende Zahl der Äusserungen (rund 96 %) deskriptiv... und zwar grösstenteils positiv" sind (Scheerer u.a. 1985, S. 187). Wenn die Autoren von den kritisierten Verhaltensweisen ausgehen und ein "Bild des guten Schülers" entwerfen, dann ähnelt er verblüffend dem willigen und angepassten Schüler aus der alten Zensurenzeit.

Die Terminologie der Autoren zu entwicklungsbezogenen vs. normativen Aussagen ist etwas verwirrend (S. 190), aber sie stellen fest : "Aussagen zur Entwicklung finden sich zwar häufiger als solche normativen Charakters, aber insgesamt sind nur 2,4 % aller Aussagen auf die Entwicklung bezogen" (S. 190). Das spricht nicht für die Anwendung individueller Bezugsnormen.

In ihrer Zusammenfassung schreiben die Autoren : "Das wichtigste Ziel der Reform, die individuelle, an der Entwicklung des einzelnen Schülers orientierte Beschreibung und Beurteilung seines Sozial- und Arbeitsverhaltens, findet sich in den untersuchten Zeugnissen kaum verwirklicht. Nicht die Persönlichkeit des Schülers, sondern die der Lehrerin bestimmt die Auswahl und die Kombination der Inhaltskategorien, die angesprochen werden. Zwar differenziert jede Lehrerin zwischen ihren Schülern, aber jeweils nur im Rahmen der ihr wichtigen Ziele, deren Auswahl auch durch die Persönlichkeitsvariablen wie Dienstalter und Geschlecht mitbestimmt sind ... Nur ein Vergleich mit anderen Zeugnissen derselben Lehrerin versetzt Eltern in die Lage zu sehen, was ihr wichtig ist. Noch problematischer wird es, wenn man berücksichtigt, dass die Eltern über die Problematik des Sozialverhaltens, aber auch des Arbeitsverhaltens nichts wissen" (Scheerer u.a. 1985, S. 194 f.). Die Autoren betonen noch, dass sie nicht überprüfen konnten, wie sich der Wegfall der Ziffernzensur auf Selbstbewusstsein der Schüler und Klassenklima auswirkte. Und natürlich waren auch hier keine Aussagen über die Validität der Zeugnisse möglich.

Die Querschnittuntersuchung von Benner und Ramseger (1985) ist schwer zu beurteilen, da sie nicht den üblichen empirischen Forschungsmethoden folgt. Die Autoren haben u.a. eine formale und eine inhaltliche Auswertung von ca.

450 Grundschulzeugnissen vorgenommen. Sie geben keine Hinweise zur Stichprobenauswahl und beanspruchen keine Repräsentativität. Es gelang ihnen aber ihrer Meinung nach, "eine Typisierung der Zeugnisse nach pädagogisch-systematischen Gesichtspunkten vorzunehmen, in welcher sich die Lehrer, die die Zeugnisse geschrieben hatten, nach Rückkoppelung der Ergebnisse durchaus wiedererkannten" (S. 154). Wieviel Lehrer dazu befragt und wieviel sich zu welchem Grade wiedererkannten, wird allerdings nirgendwo gesagt. Die Autoren kommen zu vier Idealtypen, dem normativen, dem schönen, dem deskriptiven und dem entwicklungsbezogenen Zeugnis. Beim normativen Zeugnis wird der Beurteilte an unabhängig von seiner "Person feststehenden Leistungs- und Verhaltensstandards gemessen" (S. 157), an deren Bestimmung er nicht selbsttätig mitwirkt. Dieses Zeugnis kann einen hohen Informationswert haben, aber es hilft kaum weiter, "weil die in ihm erteilten Ratschläge tautologisch sind" (S. 158).

Auch das "schöne Zeugnis" soll normative Züge haben, sei jedoch eher durch eine pluralistische Normativität gekennzeichnet. Es suche bei jedem Kind etwas Positives und sei bemüht, fehlende Leistungen positiv zu umschreiben (S. 159). Ohne Zweifel sind die Lehrer den Empfehlungen zur "ermutigenden Erziehung" gefolgt, aber das Ergebnis ist nach Meinung der Autoren ein Zeugnistyp, der "bei auch äusserlich geringstem Umfang nur minimale Informationen beinhaltet".

Aus den Beispielen von Benner und Ramseger (1985, S.159) werden die "Hinweise zu den Lernbereichen" aus einem "schönen" Zeugnis angeführt : "Auch im Lesen und Schreiben macht Peter in jüngster Zeit Fortschritte. Bekannte Texte liest er fliessend und gut betont. Es fällt ihm auch nicht mehr so schwer, unbekannte Texte zu lesen. Geübte Texte kann er fehlerfrei aus der Vorstellung aufschreiben. Es fällt ihm aber noch schwer, bei ungeübten Sätzen Lösungshilfen einzusetzen. Peter beherrscht im wesentlichen den Zahlenraum 1-100 mit allen vier Grundrechenarten und kann auch leichte Sachaufgaben selbständig lösen. Im Sachunterricht arbeitet Peter interessiert mit."

Das "deskriptive Zeugnis" verzichte auf beschönigende Tendenzen und bestimme die Leistungsnormen aus der Entwicklung des Schülers selbst. Die Empfehlungen seien durchaus pragmatisch. "Das deskriptive Zeugnis versucht, ohne zu beschönigen oder zu verurteilen, lediglich zu schildern, was war" (S. 165). Da die Autoren an keiner Stelle zu erkennen geben, woher sie wissen, was in der Klasse geschah, müssen wir einschränken, dass die Formulierungen bei ihnen den Eindruck erweckten, es handele sich um Tatsachenbeschreibungen. Aber auch dieses Zeugnis ist nach Meinung der Autoren pädagogisch defizitär.

Positiv sehen die Autoren den vierten Typ, "das entwicklungsbezogene Zeugnis". Es schildere Prozesse, bemühe sich, die wechselseitige Abhängigkeit von Arbeits-, Sozial- und Leistungsverhalten durch methodische und thematische Verknüpfung im Blick zu behalten, es beschreibe Lernsituationen als Lebenssituationen und verweise damit auf die Verantwortung der Schule, solche Situationen auch bereitzustellen (S. 165). Die Autoren meinen : "Das entwicklungsbezogene Zeugnis spiegelt einen Unterricht wider, der von dieser Verpflichtung zum persönlichen Engagement für das individuelle Kind geprägt ist. Er ermutigt durch die Beschreibung einer ermutigenden Praxis" (S. 166). Da über den Unterricht so wenig bekannt ist wie über die Wirkungen dieser Zeugnisse, ist das reine Vermutung wie alles, was die Autoren sonst noch über den Unterricht der Zeugnisschreiber sagen.

Auch für diesen Typ soll ein Beispiel mit "Hinweisen zu den Lernbereichen" angeführt werden : "Sein anfängliches Ausweichverhalten baute er langsam ab, als er kleine Sätze erlesen konnte und seine Lesefähigkeit im Spiel und bei der Klärung von Sachfragen anwenden lernte. Er ist jetzt in der Lage, kurze zusammenhängende Texte zu erlesen und zu verstehen. Frank entwickelte seinen Schreibstil langsam, vorsichtig und manchmal etwas verkrampft. Bald aber konnte er Buchstaben formen, und Wortverbindungen gelangen ihm fast fehlerfrei. Die an ihn gestellten Anforderungen bewältigte er ohne nennenswerte Hilfe. Frank arbeitete konzentriert und liess sich selten ablenken. Obwohl Frank Plus- und Minusgleichungen im Zahlenraum bis 10 lösen, Vergleiche mit entsprechenden Zeichen darstellen und Mengenoperationen richtig durchführen kann, brauchte er häufig Zuspruch von Erwachsenen, um gestellte Aufgaben zu beginnen und zügig durchzuführen. Franks grosse Stärke ist sein Informationsvorsprung in allen Bereichen des Sachunterrichts, seine Gründlichkeit, Zusammenhänge zu erkunden, und seine Neugier, mit der er seine Umwelt betrachtet. Über sein Sachinteresse sollte noch bewusster versucht werden, ihm die sogenannten Kulturtechniken anzubieten als einen Weg, weitere Informationen einzuholen" (S. 166).

Benner und Ramseger haben nicht überprüft, ob andere auch die Zeugnisse diesen vier Typen zuordnen würden. Nur in einer Anmerkung teilen sie mit, dass das "schöne" Zeugnis "mit Abstand am häufigsten" war, gefolgt vom "normativen" Zeugnis und vom seltenen "deskriptiven" Typ. Am seltensten war das "entwicklungsbezogene" Zeugnis. Genauere Angaben fehlen. Für die Beurteilung des Reformerfolges hat die Untersuchung wenig Aussagekraft.

In Niedersachsen wurden 1977 zensurenfreie Zeugnisse in den ersten beiden Grundschulklassen eingeführt. Im 3. und 4. Schuljahr sollten die Zensuren durch Verbalbeschreibungen des Arbeitsverhaltens und der besonderen

Interessen und Fähigkeiten ergänzt werden. 1979 hat Schmidt (1980, 1981) eine Untersuchung durchgeführt, in der 237 Grundschullehrer(innen), 2 653 Mütter und Väter schriftlich befragt und 1 423 Zeugnisse (902 aus der 1. und 2. Kl., 521 aus der 3. und 4. Kl.) durchgesehen wurden. Schmidt arbeitet mit grösseren, aber auch nicht repräsentativeren Zahlen als R. Schwarzer und Scheerer u.a., aber auch mit gröberen Methoden und ohne Prüfstatistik.

Im Gegensatz zu sonstigen Umfragen würden in Niedersachsen mehr als zwei Drittel der Lehrer die Reform wenigstens teilweise rückgängig machen. Das hängt sicher damit zusammen, dass die Reform hier weniger durch Tagungen und Konferenzen vorbereitet war als in NRW und dass sich die Lehrer überfordert fühlten. Die Eltern haben sich häufiger als die Lehrer, aber mit 35,7 % deutlich geringer als in NRW für die Zeugnisreform ausgesprochen. Mehr als die Hälfte der Eltern - leider ohne Aufgliederung nach Klassenstufen - wünscht sich Ziffernzeugnisse mit Kommentaren.

Die Zeugnisanalyse zeigt im 1. und 2. Schuljahr eine enge Anlehnung an die Formulierungsvorschläge des Ministeriums. Die Angaben zu "Interessen, Fähigkeiten, Fertigkeiten" wurden von den Lehrern besonders selten angemessen beschrieben. Die Lernstandsbeschreibungen waren ohne Fachkenntnisse kaum verständlich.

Da das 3. und 4. Schuljahr kaum ein anderes Bild ergaben und hier nicht so bedeutsam sind, soll gleich auf Schmidts Folgerungen eingegangen werden. Schmidt (1981, S.492) bezweifelt, "dass die Umstellung auf Verbalzeugnisse mit vorgegebenen Kategorien und Formulierungsmustern befriedigend und sachangemessen ist". Er spricht den Zeugnissen die Berichtsfunktion ab, denn "die Eltern wissen ohnehin viel mehr über das Kind, als im Zeugnis steht", und die Lehrerinnen würden sich mehrheitlich lieber ein eigenes Bild machen. Zeugnisse haben nach Schmidt auch keine Orientierungsfunktion, denn "wo gibt es den Schüler, der nicht auch ohne formelles Zeugnis orientiert wäre über seinen Wissensstand in der Klasse?" (S. 492). Die pädagogische Funktion verneint er auch, denn "es wäre eine armselige Pädagogik, die auf das Zeugnis als Medium, als Erziehungsmittel, angewiesen wäre" (1980, S. 112 und 1981, S. 493). Da auch die Berechtigungsfunktion entfiele, wenn alle Kinder gemeinsam in eine Orientierungsstufe gingen, fragt Schmidt (1981, S. 493) : "Wem würde es fehlen, wenn es in der Grundschule keine Zeugnisse gäbe ?".

Er gibt die Antwort, dass der Bürokratie eine Akte zur Legitimation von Versetzungsentscheidungen fehlen würde. Schmidt (1980, S. 114 und 1981, S. 494) schlägt vor, Zeugnisse so informationsarm wie möglich zu machen,

umfassende Information der Eltern in Elterngesprächen für verbindlich zu erklären und notfalls Briefe zu schreiben.

Diese Antworten sind angesichts der Doppelfunktion von Zensuren zu kurzschlüssig. Aber sie deuten an, in welchem Dilemma die Befürworter der Zeugnisberichte stecken, die sich auf ihre pädagogisch ehrenwerten Ziele verlassen und in 10 Jahren weder vernünftige Untersuchungen vorlegen noch ihre Praxis wissenschaftlich fundieren konnten.

Das Dilemma der Zeugnisberichte

Mit den Zeugnisberichten sollte eine Reform der Schülerbeurteilung erfolgen, ohne dass die diagnostischen Methoden verbessert wurden. Zensuren sind pädagogisch so informationsarm und messmethodisch so ungenau, dass man ohne weiteres zensurenfreie Räume schaffen kann. Aber der entscheidende Irrtum der "Anti-Zensuren-Bewegung" liegt darin, das man meint, mit dem Ersatz der Ziffer durch Verbalberichte sei auch schon eine didaktische, pädagogische und diagnostische Verbesserung verbunden. Eine Änderung der Mitteilungsform bedeutet für die Güte des diagnostischen Prozesses überhaupt nichts und erst recht nichts für die Qualität des Unterrichts. Es soll nicht unterstellt werden, dass die sich ausbreitende Praxis der Kennzeichnung von Schülerarbeiten mit Sternen, Blumen usw. anstelle der Ziffernzensur im Sinne der Befürworter von Verbalberichten ist. Aber sie haben durch ihren Verzicht auf ein diagnostisches Programm viel zu den Fehlentwicklungen beigetragen. Wenn Bolscho (1977, S. 98) sagt, dass "das Objektivitätsproblem nicht das Problem der neuen Beurteilungsformen" sei, dann distanziert er sich von wissenschaftlich fundierter Diagnostik. Diagnostik hat immer mit dem Objektivitätsproblem zu tun, denn ohne die Beschäftigung damit könnte sie nicht die Validität ihrer Aussagen erhöhen. Pädagogische Diagnostik steht wie jede handlungsbegleitende Diagnostik vor der Schwierigkeit, dass das pädagogische Handeln häufig keine optimalen Voraussetzungen zur Sicherung der Objektivität erlaubt. Wenn der Lehrer sich z.B. während des Mathematikunterrichts durch kurze Rückfragen überzeugen will, ob bestimmte Schüler seine Erklärung so gut verstanden haben, dass er weitergehen kann, dann wäre es unsinnig, auf vergleichbaren Schwierigkeitsgrad der Fragestellungen, auf gleiche Antwortbedingungen usw. zu achten. Diese diagnostischen Kurzeinschätzungen müssen und können mit messmethodischen Kompromissen leben (Ingenkamp, 1985a).

Ein weiteres Indiz dafür, dass bei der Einführung der Zeugnisberichte über den diagnostischen Prozess nicht gründlich nachgedacht wurde, ist die ständige Diskriminierung des sozialen Vergleichs. Auch wenn der Lehrer nach dem Forschungsstand z.B. keine genaue Information über die Abfolge der Konsonanten im Lauterwerb hat und daher nur höchst unvollkommen kriterienorientierte Bezugsnormen anwenden kann, ist es ihm doch möglich, beim Vergleich mehrerer Kinder zu sagen, wer in der Lautbildung mehr und wer weniger fortgeschritten sei. Immer dann, wenn das Sachgebiet noch nicht genügend durchstrukturiert ist oder wenn wir keine Zeit haben, hierarchisch strukturierte Kriterienraster anzulegen, greifen wir auf soziale Vergleiche zurück. Eine völlig andere Frage ist es, ob das Ergebnis sozialer Vergleiche dem Betroffenen so mitgeteilt wird. Sozialer Vergleich als Mitteilungsform und sozialer Vergleich als elementare Stufe im diagnostischen Vergleich sind also ganz verschiedene Dinge. Die Befürworter der Verbalberichte haben aber einerseits jeden sozialen Vergleich verteufelt und andererseits den Lehrer geradezu zu sozialen Vergleichen als elementarer Form diagnostischen Vorgehens gezwungen, weil sie ihm keine praktikablen diagnostischen Vorgehensweisen für intraindividuelle und kriterienorientierte Vergleiche an die Hand gaben. Dies ist ein zum Scheitern verurteilter Selbstbetrug.

Wer nach praktikablen diagnostischen Instrumentarien für die Eingangsstufe sucht, wird vergeblich suchen. Wo sind Beobachtungshilfen zur Sprachentwicklung, wo sind erprobte Texte, die den Prozess des Lesen- und Schreibenlernens abzutesten gestatten, wo sind Standardsituationen, wie z.B. Rollenspiele, in denen Sozialverhalten beobachtet werden könnte ? Die Liste liesse sich fortsetzen. Alles überlässt man wieder der Eigenarbeit des Lehrers. Auch die Zurückhaltung einiger Ministerien, im Sozialbereich Anleitungen zu geben, hängt ja nur vordergründig damit zusammen, dass man der Kreativität des Lehrers nicht im Wege stehen will. In Wirklichkeit will man in diesem politisch brisanten Feld keine Angriffsflächen bieten.

Durch Verordnungen, Beschwörungen und Rezeptliteratur lässt sich noch keine Reform der Schülerbeurteilung erreichen. Da die begleitende wissenschaftliche Fundierung und die Erarbeitung praktikabler diagnostischer Hilfsmittel nicht erfolgten, wird auch der anerkennenswerten Idee des Zensurenberichtes kaum ein anderes Schicksal als dem Diagnosebogen beschieden bleiben, selbst wenn das Verbalzeugnis als Routineformel überlebt.

Für die Schülerberatung hat der Zeugnisbericht kaum Bedeutung. Hier ist Schmidt (1981) zuzustimmen, dass die Lehrer mehr wissen, als sie in das Verbalzeugnis schreiben. Es eröffnet ihnen weder neue Informationen noch systematisiert es vorhandene. Das Verbalzeugnis hat, wie die Unter-

suchungen zeigen, für andere Lehrer kaum Informationswert, da diese andere Massstäbe, Schwerpunkte und Termini haben.

Literatur

Arbeitskreis Grundschule : Zeugnis : Ziffern oder Gutachten ? Frankfurt/Main, 1978 (vervielf.).

BARTNITZKY, H. und CHRISTIANI, R. : Zeugnisschreiben in der Grundschule. Düsseldorf : Bagel, 1979.

BENNER, D. und RAMSEGER, J. : Zwischen Zifferzensur und pädagogischem Entwicklungsbericht : Zeugnisse ohne Noten. In : Zeitschrift für Pägagogik, 31, 1985, 151-174.

BLUMENFELD, W. : Urteil und Beurteilung. In : Archiv für die gesamte Psychologie, 3. Ergänzungsband. Leipzig : Akademische Verlagsgesellschaft, 1931.

BOLSCHO, D. : Die Zeugnisordnungen im 1. und 2. Schuljahr. In : Bolscho, D. und Schwarzer, Ch. (Hg.) : Beurteilen in der Grundschule. München : Urban und Schwarzenberg, 1979a.

BOLSCHO, D. : Beurteilen - ein Abriss zum Stand der wissenschaftlichen Forschung. In : Bolscho, D. ; Burk, K. & Haarmann, D. : Grundschule ohne Noten. Frankfurt : Arbeitskreis Grundschule, 1979.

BURK, K. : Ziele und Inhalte der neuen Zeugnisbestimmungen - zum gegenwärtigen Stand in den Bundesländern. In : Bolscho, D. ; Burk, K. & Haarmann, D. : Grundschule ohne Noten. Frankfurt : Arbeitskreis Grundschule, 1979b.

CHRISTIANI, R. : Auswirkungen des Überganges von der zensurenbezogenen zur zensurenfreien Leistungsbewertung in den Klassen 1 und 2 der Grundschulen des Landes Nordrhein-Westfalen. In : Bund-Länder-Kommission : Evaluation in Innovationen im Bereich der Grundschule/Primarschule. Bern : Haupt, 1983.

FIRNKES, H. : Gute Noten für keine Noten ! In : Lehrer-Journal, 1984, 241-244.

INGENKAMP, K. (Hg.) : Die Fragwürdigkeit der Zensurengebung. Weinheim : Beltz, 1971, 7, 1977.

INGENKAMP, K. : Testkritik ohne Alternative. Eine kritische Darstellung der Argumentation radikaler Schultestkritik in der deutschen Fachliteratur. In : Jäger, R.S. u.a. (Hg.) : Tests und Trends, 1981. Weinheim : Beltz, 1981.

INGENKAMP, K. : Lehrbuch der Pädagogischen Diagnostik. Weinheim : Beltz, 1985a.
INGENKAMP, K. : Drei Dekaden Pädagogischer Diagnostik in Deutschland. Entwicklungen, Kontroversen und Perspektiven. In : Jäger, R.S. u.a. (Hg.) : Tests und Trends 4. Weinheim : Beltz, 1985.
LISSMANN, U. : Leistungsrückmeldungen in der Schule : Formen und Wirkungen. In : Bessoth, R. (Hg.) : Schulleitung 4, Lerneinheit 45.05. Neuwied : Luchterhand, 1984, 1-30.
MÜSSENER, U. : Leistungserziehung und Leistungsbewertung. In : Wittenbruch, W. : Das pädagogische Profil der Grundschule. Heinsberg : Agentur Dieck, 1984.
PETILLON, H. : Soziale Beziehungen zwischen Lehrern, Schülern und Schülergruppen. Weinheim : Beltz, 1982.
SCHEERER, H. ; SCHMIED, D. & TARNAI, Ch. : Verbalbeurteilung in der Grundschule. In : Zeitschrift für Pädagogik, 31, 1985, 175-200.
SCHMACK, E. : Gundula Niemands erstes Zeugnis. Kastellaun : Henn, 1977.
SCHMACK, E. : Zur neuen Schülerbeurteilung in der Grundschule. In : Pädagogische Rundschau, 32, 1978, 233-253.
SCHMIDT, H.J. : Grundschulzeugnisse in Niedersachsen. Lüneburg : Neubauer, 1980.
SCHMIDT, H.J. : Grundschulzeugnisse unter der Lupe. In : Die Deutsche Schule, 1981, 486-496.
SCHMIDT, G. & SILKENBEUMER, E. : Zensurenlose Zeugnisse. Zur Produktion abweichenden Verhaltens in der Grundschule. In : Kury, H. und Lerchenmüller, H. (Hg.) : Schule, psychische Probleme und sozialabweichendes Verhalten... Köln : Heymanns, 1983.
SCHWARZER, R. : Fallstudie zur Evaluation der Neuregelung von Zeugnissen in der Grundschule. RWTH Aachen, 1980.
SEYFRIED, H. : Zeugnisnoten. Ergebnisse und Interpretationen empirischer Untersuchungen. Wien : Ketterl, 1983.
ULLRICH, H. und WÖBCKE, M. : Notenelend in der Grundschule. München : Kösel, 1981.

Résumé

Carnets verbaux à l'école primaire en Allemagne Fédérale Objectifs, Introduction, Evaluation

Entre 1970 et 1977, les Länder de la République Fédérale Allemande ont remplacé les notes en chiffres par des rapports pour le premier et en partie pour

le deuxième degré primaire. Pour aider les enseignants à écrire ces rapports, des exemples de textes, des listes de critères et des cours complémentaires leur ont été proposés. La compétence diagnostique - notamment à l'observation - de l'enseignant n'a pas fait l'objet d'une formation. Les effets attendus de l'introduction de ces rapports étaient une diminution de l'anxiété des élèves et une augmentation de leur motivation. Certains auteurs attestent de tels effets, tandis que d'autres les mettent en doute. Très peu d'évaluations ont été effectuées ; les quelques études disponibles - dont on peut cependant douter des qualités méthodologiques - confirment l'impression que cette réforme ordonnée par l'administration n'a atteint son objectif ni en ce qui concerne l'appréciation du travail ni par rapport à l'amélioration de l'enseignement.

Summary

Verbal reports in elementary school in Germany : Goals, introduction, evaluation

The Laender of the Federal Republic of Germany changed their reporting system from 1970-1971. Pupils in classes 1 and sometimes 2 of the elementary school receive now essay type reports instead of numeric marks. Teachers got help for the wording of the reports by model reports, criterion lists or training. But their diagnostic competence especially in observing pupils has not been trained. Some believed the new reports would lower anxiety and encourage motivation. Others doubted these expectancies. Only a few studies evaluated the efficiency of the new reports, and these were weak as far as methods and design are concerned. But even with these restrictions there is more evidence that this reform did not reach its goals, neither in the field of pupils' evaluation nor with respect to the improvement of instruction.