

Steltmann, Klaus

## Der Bonner Lehrerbewertungsbogen (BLBB). Ein Fragebogen zur Beurteilung von Lehrer durch Schüler

*Unterrichtswissenschaft 14 (1986) 4, S. 402-415*



Quellenangabe/ Reference:

Steltmann, Klaus: Der Bonner Lehrerbewertungsbogen (BLBB). Ein Fragebogen zur Beurteilung von Lehrer durch Schüler - In: Unterrichtswissenschaft 14 (1986) 4, S. 402-415 - URN: urn:nbn:de:0111-pedocs-296127 - DOI: 10.25656/01:29612

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-296127>

<https://doi.org/10.25656/01:29612>

in Kooperation mit / in cooperation with:

# BELTZ JUVENTA

<http://www.juventa.de>

### Nutzungsbedingungen

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

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

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

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

### Kontakt / Contact:

**pedocs**  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Digitalisiert

Mitglied der

  
Leibniz-Gemeinschaft

## Der Bonner Lehrerbewertungsbogen (BLBB)

### Ein Fragebogen zur Beurteilung von Lehrern durch Schüler

In diesem Beitrag wird ein Instrument zur Erfassung des Lehrerverhaltens durch Schüler vorgestellt. Es enthält 36 Items mit den Dimensionen:

- emotionale Einschätzung
- Unterrichts-Effektivität
- Anforderungsniveau.

Das Verfahren eignet sich zur Messung des Klassenklimas in empirischen Untersuchungen, aber auch zur Selbstdiagnose von Lehrern, die ihren Unterricht verbessern möchten.

#### A questionnaire for teacher evaluation by pupils

In this article an instrument for the evaluation of teaching profiles with the help of students is presented. It contains 36 items, which fall into three categories:

- emotional climate in the classroom
- instructional efficiency
- standards of the teacher.

The questionnaire can be employed for measuring classroom climate in empirical studies as well as providing teachers with an instrument of self-analysis.

## 1. Problemstellung

### 1.1 Inhaltliche Fragen

Inhaltlich ordnet sich die vorliegende Untersuchung in die Bemühungen um die Erforschung des Lehrerverhaltens ein. In Deutschland stammen die bedeutendsten Arbeiten zu diesem Thema von *Reinhard* und *Annemarie Tausch*, die in der letzten Auflage ihrer „Erziehungspsychologie“ (Göttingen 1977) fünf Dimensionen des Erziehverhaltens auführen:

- Dirigierung, Lenkung
- Achtung – Wärme – Rücksichtnahme (emotionale Dimension)
- Einführendes, nicht-wertendes Verstehen
- Echtheit
- Fördernde, nicht-dirigierende Einzeltätigkeiten.

Von diesen Dimensionen sind die letzten drei empirisch wenig abgesichert, da sie vom Ehepaar *Tausch* in späteren Auflagen zugefügt wurden, als sie sich von empirischen Methoden fort stärker einer ganzheitlich-humanistischen Betrachtungsweise zugewandt hatten. In der „Erziehungspsychologie“ fällt auch auf, daß keinerlei Unterschied zwischen Erziehverhalten in verschiedenen Situationen (z.B. Schule/Familie) oder in verschiedenen Altersstufen gemacht wird. Eine sinnvolle Erfassung des Lehrerverhaltens muß jedoch meines Erachtens auch die spezifisch unterrichtliche (didaktische) Komponente erfassen.

Eine spezifischer auf das Lehrerverhalten bezogene Untersuchung war die von

Ryans (Characteristics of Teachers, 1964), bei der mehrere tausend Lehrer mit Hilfe von Schätzskaletn durch Unterrichts-Beobachter eingeschätzt wurden. Ryans' Untersuchung ergab drei Dimensionen des Lehrerverhaltens:

- warmes, verstehendes vs. distanziertes, egozentrisches Unterrichtsverhalten (emotionale Dimension)
- sachliches, systematisches vs. planloses, nachlässiges Verhalten
- stimulierendes, anregendes vs. langweiliges, routiniertes Verhalten.

Ryans' Theorie des Lehrerverhaltens bietet gute Anknüpfungspunkte für eine neue Untersuchung mit Hilfe von Schätzskaletn für Schüler; ich komme unter Punkt 3 noch einmal darauf zurück.

## 1.2 Methodische Fragen

Eine der wichtigsten Aufgaben der Unterrichtsforschung ist es, verlässliche Verfahren zur Erfassung des Unterrichts-Geschehens zu entwickeln. Das Problem, die Ganzheit des Unterrichts analytisch in den Griff zu bekommen, ist komplex und wird von keinem der bisher entwickelten Verfahren befriedigend gelöst. Beobachtungsverfahren, die mit Beobachtungskategorien arbeiten (z.B. *Flanders*, *VICS*), sind sehr aufwendig im Einsatz, haben immer Schwierigkeiten mit der Abgrenzung der Kategorien gegeneinander und erfassen stets nur einen begrenzten Unterrichtsaspekt. Schätzskaletn sind dagegen in stärkerem Maße der Subjektivität der Schätzer unterworfen, die sich auch durch eine noch so gute Schulung nicht ausschließen läßt. Vor allem ist bei allen Verfahren (Kategoriensysteme, Schätzskaletn) der Stichprobencharakter der Beobachtung eine Quelle ganz erheblicher Verzerrungen. Insbesondere bei kurzfristiger Beobachtung ist die Gefahr gegeben, daß sich Lehrerverhalten gerade durch die Beobachtung selbst verändert, und zwar in der Regel in Richtung auf sozial erwünschte Verhaltensweisen. Erst bei längerer Beobachtung setzt sich allmählich das „wahre“, alltägliche Lehrerverhalten wieder durch. Gerade eine solche längere Beobachtung ist jedoch bei den meisten Forschungsvorhaben wegen der Begrenztheit von Zeit und Mitteln nicht zu realisieren.

Auf diesen Überlegungen basiert der Einsatz von Schülern als Unterrichts-Beobachtern, denn Schüler erleben den Unterricht täglich, beurteilen ihn also aufgrund einer Gesamterhebung und nicht aufgrund von zufälligen und möglicherweise verzerrten Stichproben des Lehrerverhaltens. Allerdings lassen sich Schüler nur mit Hilfe von Schätzskaletn als Unterrichts-Beobachter einsetzen, bei denen sie Eindrücke aus dem früheren Unterricht generalisieren müssen. Gegenüber Schätzskaletn für nicht-teilnehmende Beobachter haben solche für Schüler neben der Vergrößerung der Verhaltensstichprobe zur Gesamterhebung noch eine Reihe weiterer Vorteile.

Der erste Vorzug einer Schätzskaletn für Schüler liegt in der erheblichen Vergrößerung der Zahl der Beobachter. Wegen der Beobachtungsfehler (implizite Persönlichkeitstheorie, Halo-Effekt usw.), denen auch geschulte Beobachter noch unter-

liegen, werden in sorgfältigen Untersuchungen häufig zwei oder mehr Beobachter eingesetzt und die Mittelwerte ihrer Beobachtungen als Ergebnis gewertet. Einer erheblichen Vermehrung der Beobachter sind jedoch in der Regel finanzielle und organisatorische Grenzen gesetzt. Zudem ist eine Veränderung des Lehrerverhaltens aufgrund der Anwesenheit der Beobachter nicht auszuschließen; sie wird vermutlich um so größer, je mehr Beobachter eingesetzt werden. Verwendet man dagegen Schätzskalen für Schüler, so gewinnt man ohne besonderen Aufwand 10 bis 40 Beurteiler, die zudem den Unterricht nicht beeinträchtigen können.

Nun wird man einwenden können, daß Schüler zum einen kaum als Beobachter geschult werden können und daß sie zum anderen subjektiven Einschätzungen besonders stark unterworfen seien. Beides ist richtig, doch muß die Subjektivität der Schüler nicht unbedingt von Nachteil sein. Denn während die Subjektivität eines außenstehenden Beobachters nichts mit dem Unterricht zu tun hat –, also reiner Störfaktor für die Untersuchung ist, prägt die subjektive Wahrnehmung des Lehrerverhaltens durch die Schüler ganz entscheidend das gesamte Unterrichtsklima. Etwas pointiert könnte man behaupten: Nicht auf das „objektive“ Verhalten des Lehrers im Unterricht kommt es an, sondern darauf, wie es subjektiv von den Schülern wahrgenommen wird. Nicht ob ein Lehrer nach den Maßstäben eines „objektiven“ Unterrichtsbeobachters eine „zwanglose Unterrichtsatmosphäre“ schafft, ist entscheidend für das Verhältnis zwischen Lehrer und Klasse, sondern wie die Unterrichtsatmosphäre von der Klasse selbst empfunden wird. Insofern stellt die Subjektivität der Schüler keine Beeinträchtigung oder Verzerrung der Meßgenauigkeit einer Skala dar, sondern trägt geradezu zur Validität des Verfahrens bei.

Es muß allerdings deutlich gesehen werden, daß Schüler den gleichen psychologischen Beurteilungsfehlern unterliegen wie andere Beurteiler auch. Mit dem Halo-Effekt, mit der Tendenz zur Mitte, mit dem Milde-Fehler, vor allem auch mit der Tendenz zur Beantwortung der Fragen in Richtung auf sozial erwünschtes Verhalten muß auch bei Schätzskalen für Schüler gerechnet werden. Letzteres, zusammen mit dem Halo-Effekt, bewirkt, daß ein beliebter Lehrer wesentlich günstiger beurteilt wird als ein unbeliebter.

Eine starke Abhängigkeit der Beurteilung von der Stimmungslage der Klasse ist ebenfalls nicht zu verkennen. Hat die Klasse sich gerade über eine Entscheidung des Lehrers geärgert, wird sie es ihm möglicherweise durch eine ungünstigere Beurteilung „heimzahlen“. Ein vollkommen problemloses Instrument zur Erfassung des Unterrichts ist also auch eine Schätzskala für Schüler keineswegs.

Trotzdem stellen Schätzskalen für Schüler insgesamt gesehen eine gute und besonders auch ökonomische Möglichkeit zur Einschätzung des Lehrerverhaltens bzw. des Unterrichtsklimas dar. Umfangreiche Forschungen in den USA – allerdings fast ausschließlich auf College-Niveau – haben gezeigt, daß diese Instrumente durchaus valide und reliabel sein können (vgl. dazu *Marsh u.a., 1975*).

## 2. Forschungsstand

In Deutschland liegt bisher ein Instrument zur Erfassung des Lehrerverhaltens durch Schüler vor: die Dortmunder Skala zur Lehrerbeurteilung durch Schüler (DSL). Die Autoren der DSL gehen von einem Zwei-Faktoren-Modell elterlicher Erziehungsstile aus, das von Herrmann, Schwijatewski und Ahrens (1968), Herrmann, Stapf und Krohne (1971) sowie Stapf, Herrmann und Stäcker (1972) entwickelt wurde. Es enthält die beiden Komponenten „Elterliche Strenge“ und „Elterliche Unterstützung“. Die Autoren übertragen dieses Modell auf die Schulsituation, vereinigen aufgrund methodenkritischer Überlegungen die beiden Faktoren allerdings in einer gemeinsamen bipolaren Dimension, in der „Strenge“ und „Unterstützung“ die beiden Pole einer Skala bilden.

Die Beschreibung dieser Dimension im Testmanual sowie die 25 Items der Skala erwecken den Eindruck, daß die Autoren hier zwei Dimensionen zusammenfassen, die in der bisherigen Unterrichtsforschung meist getrennt als „Lenkung-Kontrolle“ mit den Polen stark/schwach und eine „emotionale Dimension“ mit den Polen Zuwendung/Ablehnung aufgeführt werden. Ob nicht auch in den Items der DSL zwei verschiedene Faktoren stecken, müßte durch eine Faktorenanalyse geklärt werden; die Autoren machen leider keine Angabe darüber, ob eine solche durchgeführt wurde.

Die DSL enthält 25 Items, die in den Stufen „immer“ („oft“), „manchmal“, „selten“ und „nie“ beantwortet werden können. Die Zusammensetzung der Eichstichprobe ist leider nicht befriedigend: Klassen aus Schulen für Lernbehinderte sind deutlich überrepräsentiert (55 von 95 Klassen), Klasse 10 völlig unterrepräsentiert (1 Klasse). Da Unterschiede in der Beantwortung der Items sowohl in verschiedenen Schultypen wie in verschiedenen Klassenstufen zu vermuten sind, wäre eine neue Eichung der DSL mit einer repräsentativen Stichprobe dringend erforderlich.

## 3. Zur Entwicklung des Bonner Lehrerbewertungsbogens (BLBB)

Wie die Ergebnisse von Ryans deutlich zeigen (s. 1.1), ist die Erfassung der Dimension „Strenge/Unterstützung“ im Unterricht sicherlich nicht ausreichend. Es müssen auf jeden Fall Items hinzukommen, die die didaktische Komponente des Unterrichts einbeziehen. In Weiterführung der Ergebnisse von Ryans war auch zu überlegen, ob nicht das Anforderungsniveau des Lehrers eine besondere Rolle spielen könnte.

Bei der Konstruktion unseres Lehrerbewertungsbogens sind wir (die Studenten mehrerer Seminare über empirische Forschungsmethoden unter meiner Leitung) von den 18 Items des Bewertungsbogens von Ryans für nicht-teilnehmende Unterrichts-Beobachter ausgegangen. Die dort als Eigenschaften gefaßten Bewertungen wurden soweit wie möglich in Verhaltensweisen transportiert, die abstrakten Items konkretisiert, manchmal in mehrere zerlegt und schließlich aus anderen Instrumenten und der Literatur ergänzt.

Unsere ursprüngliche Sammlung von 90 Items wurde nun in vielen Diskussionen mit Schülern und Lehrern daraufhin überprüft, welche von ihnen für die Erfassung des Lehrerverhaltens besonders geeignet und welche Fragen inhaltlich von größter Bedeutung seien.

In einer ersten Fassung enthielt der BLBB schließlich 42 Items, die an einer Stichprobe von 16 Klassen mit ca. 400 Schülern überprüft wurden. Als Zielgruppe für den BLBB hatten wir die Klassen 5–13 an allgemeinbildenden Schulen mit Ausnahme von Sonderschulen vorgesehen. Bei der ersten Erprobung wurde noch einmal besonders darauf geachtet, ob alle Items in allen Klassenstufen (insbesondere in 5 und 6) verständlich sind.

Die Items bestanden jeweils aus zwei Polen, zwischen denen auf einer 7stufigen Skala ein Urteil abgegeben werden sollte, z.B. zensiert gerecht 1 2 3 4 5 6 7 n (= weiß nicht) zensiert ungerecht.

Nach der ersten Erprobung wurden 10 Items ausgeschieden, und zwar aufgrund folgender Kriterien:

- weil sie von mehr als 10 Schülern nicht beantwortet wurden (Nr. 32 und 42)
- weil die Streuung innerhalb der Klasse zu groß war ( $s > 1.5$ ; Nr. 3, 7, 39 und 41)
- weil sie zu wenig zwischen verschiedenen Klassen differenzierten ( $ETA < 0.3$ ; Nr. 7 und 41)
- wegen zu geringer Ladung auf einem der Hauptfaktoren ( $Ladung < 0.3$ ; Nr. 2, 11, 16 und 39)
- wegen zu geringer Split-Half-Reliabilität ( $r < 0.8$ ; Nr. 5, 7, 11, 14, 28, 32, 40, 41 und 42)\*

Wie man sieht, konnten die meisten der ausgeschiedenen Items mehreren Kriterien nicht genügen.

In der zweiten Version wurden 4 Items neu eingefügt, die die 10 gestrichenen inhaltlich ersetzen sollten; außerdem wurden die in Version 1 unterschiedlich gepolten Items so geordnet, daß das positive Verhalten jeweils auf der linken Seite der Skala erschien, weil die unterschiedliche Anordnung Verwirrung bei den Schülern ausgelöst hatte. Die Beantwortungsmöglichkeit „weiß nicht“ wurde gestrichen, um möglichst vollständig ausgefüllte Bewertungsbögen zu erhalten.

Die Erprobung der zweiten Version (in 15 Klassen mit 380 Schülern) ergab gute Werte für die 4 neuen Items, so daß alle 36 in die dritte und endgültige Version übernommen werden konnten. Die Zahlen zwischen den beiden Polen der Items wurden jedoch durch neutrale leere Kästchen ersetzt, da die Zahlen häufig – obwohl sie von 1 bis 7 reichten – mit Noten gleichgesetzt worden waren. Die endgültige Version 3 des BLBB ist im Anhang wiedergegeben.

---

\* Dabei wurden die Klassen nach Zufall in 2 Hälften geteilt und die Mittelwerte der Klassenhälften A mit denen von B korreliert.

## 4. Eigenschaften des BLBB

### 4.1 Faktorenstruktur

Mehrere Faktorenanalysen der Daten aus unterschiedlichen Stichproben ergaben jeweils 3 Faktoren, die in jeder Stichprobe identisch blieben. Diese drei Faktoren sind:

- I. die emotionale Qualität des Unterrichts
  - II. die Effektivität des Unterrichts
  - III. das Anforderungsniveau des Unterrichts.
- (Zur genauen Analyse siehe Tabelle 6.)

Markiervariablen für Faktor I – die emotionale Dimension – sind folgende Items (Ladungen in Klammern):

- schafft eine zwanglose Atmosphäre (0.90)
- zeigt Verständnis für die Probleme der Schüler (0.87)
- ist beliebt (0.87)
- ist freundlich (0.87)
- führt den Unterricht locker (0.87)
- läßt Schülermeinungen gelten (0.85)
- nimmt die Schüler ernst (0.84)
- man kann ihm vertrauen (0.83)
- ist mir sympathisch (0.83)
- ermutigt die Schüler (0.80)

Markiervariablen für Faktor II – Unterrichtseffektivität – sind die Items:

- man lernt viel bei ihm (0.92)
- hält gute Disziplin (0.86)
- sein Unterrichtsziel ist klar zu erkennen (0.84)
- sein Unterricht ist systematisch (0.84)
- beherrscht sein Fach gut (0.82)
- kommt immer gut vorbereitet in den Unterricht (0.74)
- drückt sich verständlich aus (0.74)
- sein Unterricht ist interessant (0.73)
- kann Probleme gut erklären (0.68).

Faktor III – Anforderungsniveau des Unterrichts – wird durch hohe Ladungen auf den folgenden Variablen charakterisiert:

- stellt geringe Anforderungen (0.69)
- zensiert milde (0.64)
- gibt wenig Hausaufgaben auf (0.62)
- bei ihm kommen alle mit (0.55).

Faktor I erklärt 35.2% der Gesamtvarianz des Fragebogens, Faktor II 25.5%, Faktor III 7.7%, alle drei zusammen 68.4%.

An der Faktorenstruktur des BLBB fällt folgendes auf:

1. Die emotionale Dimension hat mit Abstand die größte Bedeutung (18 von 36 Items gehören zu ihr; zusätzlich haben 6 weitere Items hohe Ladungen ( $> 0.4$ ) in der emotionalen Dimension: Nr. 2, 10, 14, 18, 21 und 22).
2. Ryans' Dimensionen 2 und 3 (sachliches, systematisches Verhalten; stimulierendes, anregendes Verhalten) erscheinen im BLBB nur als *eine* Dimension. Item 10 (sein Unterricht ist interessant), das wir als Markiervariable für „stimulierenden Unterricht“ erwartet hatten, und andere Items (Flexibilität, Humor)

erscheinen mit Item 25 (man lernt viel bei ihm) in der gemeinsamen Dimension II („Unterrichtseffektivität“), in die insgesamt 14 Items einzuordnen sind.

- Das Anforderungsniveau des Lehrers stellt sich – wie erwartet – als eigene Dimension heraus (4 Items).

#### 4.2 Reliabilität

Die Reliabilität der einzelnen Dimensionen wurde als Split-Half-Reliabilität nach der Formel von *Rulon* (*Lienert*, 1969, Formel 58 a, S. 222) berechnet, indem jede Dimension in zwei Hälften geteilt wurde. Die errechneten Werte sind erstaunlich hoch, nämlich

- für Dimension 1:  $r_{tt} = 0.987$
  - für Dimension 2:  $r_{tt} = 0.961$
  - für Dimension 3:  $r_{tt} = 0.924$  (mit nur 4 Items!)
- für den Gesamtwert  $r_{tt} = 0.987$ .

#### 4.3 Eichung

Die Eichung wurde an insgesamt 502 Klassen vorgenommen. Der Schwerpunkt liegt im Bereich des Gymnasiums (224 Klassen), aber auch die Realschule (mit 113 Klassen) und die Hauptschule (mit 126 Klassen) sind noch genügend repräsentiert. Die Gesamtschule ging mit 39 Klassen in die Stichprobe ein. Die Verteilung der biographischen Daten entspricht annähernd den Bedingungen einer repräsentativen Stichprobe (s. Tabellen 1–5).

Es kann aber trotzdem leider nicht von einer repräsentativen Stichprobe gesprochen werden, weil man bei einer Erhebung dieser Art auf die freiwillige Teilnahme

**Tabelle 1:** Geschlecht der Lehrer (N = 502)

	Stichprobe				NRW %	
	männl.	%	weibl.	%	männl.	weibl.
Hauptschule	61	48.4	65	51.6	48.2	51.8
Realschule	45	39.8	68	60.2	42.7	57.3
Gymnasium	135	60.3	89	39.7	61.4	38.6
Gesamtschule	20	51.3	19	48.7	?	?

**Tabelle 2:** Alter der Lehrer am Gymnasium (N = 224)

	f	%	NRW %
bis 30	43	19.2	16.4
31–40	94	42.0	43.3
41–50	55	24.6	25.9
51–60	29	12.9	13.0
über 60	3	1.3	1.4

**Tabelle 3:** Klassenstufen (N = 502)

	HS	RS	Gym	GS	gesamt
Klasse 5	8	9	15	6	38
Klasse 6	15	13	15	3	46
Klasse 7	21	14	20	4	59
Klasse 8	32	25	22	5	84
Klasse 9	29	24	27	5	85
Klasse 10	21	28	23	3	75
Klasse 11	—	—	35	4	39
Klasse 12	—	—	34	4	38
Klasse 13	—	—	33	5	38

**Tabelle 4:** Fächer

	HS	RS	Gym	GS	gesamt
Religion	14	20	20	4	58
Deutsch	25	18	28	8	79
Latein	—	—	10	1	11
Englisch	19	14	25	1	59
Französisch	—	7	12	1	19
Spanisch	—	—	2	—	2
Mathematik	17	12	21	16	66
Erdkunde	7	3	10	—	20
Geschichte	7	8	12	1	28
Philosophie	—	—	5	—	5
Sozialkunde	1	1	10	1	13
Pädagogik	—	—	18	—	18
Wirtschaftskunde	4	—	—	—	4
Physik	7	3	1	1	12
Chemie	5	5	11	—	21
Biologie	6	10	16	—	32
Musik	4	1	1	—	6
Kunst	3	3	7	1	13
Sport	7	8	13	3	32
andere	—	—	3	—	3

**Tabelle 5:** Geschlecht der Schüler (N = 12 397)

	f	%	NRW%
männlich	6186	49.9	49.4
weiblich	6211	50.1	50.6

der Lehrer angewiesen ist. Diejenigen Lehrer, die sich der Teilnahme an einer Bewertung durch ihre Schüler verweigern, dürften aber gerade diejenigen sein, die eine weniger günstige Beurteilung erwarten. Die Ergebnisse werden also durchweg in Richtung auf die positive Seite verschoben sein. Aus einer schulinternen Umfrage unter Schülern, bei der eine Frage nach der Fachbeherrschung ungefähr identisch war mit dem Item „beherrscht sein Fach gut“ im BLBB, kann ich schließen,

daß die Ergebnisse bei einer nicht ausgelesenen Stichprobe um etwa 0.5 Punkte niedriger liegen könnten als bei meiner Stichprobe von freiwillig teilnehmenden Lehrern.

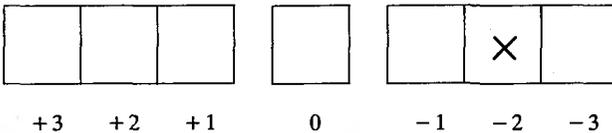
Eine weitere Beeinträchtigung der Repräsentativität liegt in der regionalen Beschränkung der Stichprobe. Die Befragung wurde von ca. 300 Studenten in mehreren Seminaren im weiteren Umkreis der Universität Bonn durchgeführt, und zwar im wesentlichen im Rheinland und im Ruhrgebiet zwischen Mainz und Dortmund. Diese regionale Begrenzung dürfte der Repräsentativität der Stichprobe jedoch weniger schaden als die Freiwilligkeit der Teilnahme.

Die statistischen Merkmale der Stichprobe können den folgenden Tabellen entnommen werden. Zum Vergleich werden – soweit zugänglich – die entsprechenden Zahlen aus der Schulstatistik des Landes Nordrhein-Westfalen angegeben (Quellen: Statistisches Jahrbuch des Landes Nordrhein-Westfalen, 1980; Jahrbuch des Philologen-Verbandes NRW 1980).

### 5. Zur Durchführung und Auswertung des BLBB

Die Durchführung des BLBB ist einfach; lediglich in 5. und 6. Klassen müssen gelegentlich einzelne Ausdrücke im Fragebogen erläutert werden. Die Durchführungszeit beträgt 20–25 Minuten. Der Lehrer sollte möglichst nicht selbst anwesend sein, um Verzerrungen in der Beantwortung zu vermeiden.

Zur Auswertung werden den Kästchen Zahlen zugeordnet, und zwar in folgender Weise.



Das im Beispiel angekreuzte Kästchen wird also als  $-2$  verrechnet. Zu jedem Item im Fragebogen wird das arithmetische Mittel aus den Antworten aller Schüler berechnet; dieses kann dann mit dem Norm-Mittelwert in Tabelle 6 verglichen werden. 68% aller Lehrer liegen im Bereich von  $\pm s$  um diesen Mittelwert, also z.B. bei Item 1 im Bereich von 0.5 bis 1.9. Bewertungen in diesem Bereich sind als durchschnittlich anzusehen; außerhalb davon sind sie als über- bzw. unterdurchschnittlich zu interpretieren.

Aus Tabelle 6 können folgende Daten entnommen werden:

- $\bar{x}$  ist das arithmetische Mittel aller Klassen
- $s$  ist die Standardabweichung der Klassenmittelwerte
- ETA (Lehrer) ist ein Maß dafür, wieviel der Varianz dieses Merkmals durch die Unterschiede zwischen den einzelnen Lehrern erklärt wird; je höher ETA, um so höhere Unterscheidungskraft hat das Item, bzw. um so stärker unterscheiden sich die Lehrer in diesem Merkmal voneinander.

ETA (Klasse) ist das gleiche Maß für Unterschiede zwischen den Klassen. Es ist nicht mit ETA (Lehrer) identisch, weil viele Lehrer mit mehreren Klassen in der Stichprobe vertreten waren; es kommen also noch die Unterschiede zwischen verschiedenen Klassen derselben Lehrer hinzu.

Ladung  
Comm. gibt die Faktorenladungen jedes Items für die Analyse aller Klassenmittelwerte an  
ist ein Maß für die gemeinsame Varianz eines Items in allen drei Faktoren.

Die Daten der Tabelle 6, insbesondere die Mittelwerte, haben nur unter der Voraussetzung der freiwilligen Teilnahme der Lehrer Gültigkeit.

**Tabelle 6:** Kennwerte des BLBB

Item	X	s	ETA (Lehrer)	ETA (Klasse)	Ladung auf Faktor			Comm.
					1	2	3	
1	1.2	0.7	28.2	36.8	.59	.50	.12	0.62
2	1.3	0.7	20.7	35.6	.41	.74	.04	0.71
3	1.3	0.9	24.8	33.3	.75	.16	.03	0.58
4	1.5	0.8	29.5	37.8	.85	.11	.06	0.73
5	1.8	0.7	29.4	38.0	-.02	.74	.40	0.71
6	0.7	1.0	42.0	49.1	.73	-.06	.04	0.53
7	2.1	0.7	17.5	24.5	.15	.82	.07	0.70
8	1.5	0.8	33.9	40.0	.87	.21	-.03	0.81
9	1.8	0.7	27.3	32.8	.87	.24	.23	0.87
10	0.7	0.9	26.3	33.2	.42	.73	-.04	0.72
11	0.9	0.7	27.4	33.4	.72	.08	.47	0.74
12	1.0	0.9	56.4	59.9	.64	.25	.33	0.56
13	1.4	0.7	17.8	22.0	.00	.84	.23	0.75
14	2.0	0.6	22.1	31.8	.48	.53	.25	0.59
15	1.2	0.8	33.2	40.8	.65	.41	-.06	0.59
16	1.1	1.0	45.8	52.2	.87	.27	-.06	0.83
17	1.3	0.7	32.2	38.0	.14	.73	-.15	0.58
18	1.6	0.7	25.5	31.3	.47	.77	-.05	0.82
19	1.4	0.7	24.2	33.0	.36	.68	-.13	0.59
20	1.3	0.9	32.6	39.5	.87	.02	-.02	0.76
21	0.4	0.8	35.9	45.6	.41	-.02	.64	0.56
22	1.0	0.8	25.1	32.5	.46	.24	.55	0.56
23	1.6	0.8	29.2	35.2	.83	.30	.10	0.78
24	1.2	0.8	27.9	37.4	.80	.34	.12	0.77
25	1.3	0.9	26.3	35.7	.14	.92	.03	0.86
26	1.3	0.8	33.2	40.5	.90	.04	.01	0.81
27	1.3	0.7	17.9	23.9	.07	.54	.27	0.36
28	1.8	0.7	35.5	36.3	.76	.26	-.20	0.69
29	0.5	0.9	25.3	31.7	-.15	.61	.51	0.65
30	1.4	0.7	24.1	27.0	.84	.22	.15	0.78
31	1.2	0.6	17.6	24.5	.18	.84	.16	0.76
32	1.1	0.9	28.9	36.1	-.04	.86	.12	0.76
33	1.1	1.1	51.0	60.0	-.17	.02	.62	0.41
34	1.3	0.7	21.5	25.2	.76	.37	.13	0.73
35	0.1	0.8	31.2	35.2	.17	.14	.69	0.50
36	1.5	0.9	31.0	33.3	.83	.33	.10	0.77
D 1	1.3	0.6	50.5	58.4				
D 2	1.4	0.6	31.9	41.1				
D 3	0.9	0.8	39.8	47.3				
Ges.	1.3	0.5	42.1	50.5				

**Tabelle 7: Durchschnittliche Mittelwertsabweichungen**

Dimension 1 (= D 1) = Items 1, 3, 4, 6, 8, 9, 11, 12, 15, 16, 20, 23, 24, 26, 28, 30, 34 und 36
Dimension 2 (= D 2) = Items 2, 5, 7, 10, 13, 14, 17-19, 25, 27, 29, 31, 32
Dimension 3 (= D 3) = Items 21, 22, 33 und 35
<b>1. Schulformen</b>
Hauptschule: D 1 -0.3, D 2 -0.1, D 3 -0.3, gesamt -0.2
Realschule: D 2 +0.1, D 3 +0.1, gesamt -
Gesamtschule: D 1 +0.2, D 2 +0.3, D 3 +0.2, gesamt +0.2
<b>2. Alter des Lehrers</b>
bis 30: D 1 +0.2, D 3 +0.1, gesamt +0.1
31-40: -
41-50: -
51-60: D 1 -0.4, D 2 -0.2, D 3 -0.4, gesamt -0.3
über 60: D 1 -0.4, D 2 -0.4, D 3 -0.6, gesamt -0.4
<b>3. Klassenstufen</b>
Klasse 5: D 1 +0.6, D 2 +0.5, D 3 +0.5, gesamt +0.5
Klasse 6: D 1 +0.4, D 2 +0.4, D 3 +0.2, gesamt +0.4
Klasse 7: -
Klasse 8: D 1 -0.1, D 2 - D 3 - gesamt -0.1
Klasse 9: D 1 -0.2, D 2 -0.1, D 3 -0.1, gesamt -0.2
Klasse 10: D 1 - D 2 -0.2, D 3 -0.1, gesamt -0.1
Klasse 11: D 1 - D 2 -0.1, D 3 -0.1, gesamt -
Klasse 12: -
Klasse 13: -
<b>4. Fächer</b>
D 1: Rel., Phil., Sozialk., Päd. +0.3; Chemie, Physik, Biol. -0.3
D 2: Mathematik +0.3, alte Sprachen und Sport +0.2, Deutsch +0.1; Erdkunde, Geschichte, Musik, Kunst -0.3
D 3: Kunst +0.8, Religion +0.6, alte Sprachen, Naturwissenschaften -0.1, neue Sprachen, Mathematik -0.2, Erdkunde, Geschichte -0.3

Will man einen Vergleich innerhalb einer bestimmten Bezugsgruppe durchführen, können Zu- und Abschläge vom Mittelwert berücksichtigt werden. Führt man z.B. eine Untersuchung in der Hauptschule durch, werden die in Tabelle 7 angegebenen Abschläge vom Norm-Mittelwert abgezogen. Für Item 1 gilt dann nicht mehr 1.2 als Mittelwert, sondern  $1.2 - 0.3 = 0.9$ . Für den einzelnen Lehrer können weitere Werte zu- oder abgezogen werden. Für einen 55jährigen Gesamtschullehrer, der eine Klasse 9 in Religion unterrichtet, ergibt sich z.B. für Item 1:  $1.2 + 0.2$  (Schulform)  $- 0.4$  (Alter)  $- 0.2$  (Klassenstufe)  $+ 0.3$  (Fach)  $= 1.1$ . 1.1 ist der Norm-Mittelwert für diesen Lehrer, an dem der tatsächlich erzielte Mittelwert gemessen wird.

### 6. Sinn und Zweck des Verfahrens

Der BLBB ist wegen seiner leichten Handhabbarkeit sicher nicht nur als Instrument empirischer Forschung anzusehen, sondern läßt sich auch informell im Einzelfall einsetzen, wo nach Ansatzpunkten für die Veränderung und Verbesserung

von Lehrerverhalten gesucht wird. Dabei ist allerdings stets zu beachten, daß Lehrerverhalten keine „objektive“ Variable ist, sondern sich in der Interaktion mit den Schülern unterschiedlich entwickelt, so daß mit dem BLBB im Grunde eher die „Klassenatmosphäre“ als lediglich das Lehrerverhalten erfaßt wird.

Insgesamt zeichnen sich für den BLBB drei Anwendungsbereiche ab:

1. bei empirischen Untersuchungen, in denen die Einschätzung des Lehrers durch die Schüler eine wichtige Variable darstellt, deren Kontrolle wünschenswert ist. Das trifft z.B. zu bei der Evaluation von Unterrichtsverfahren, beim Vergleich verschiedener Schulformen oder bei der Untersuchung der Leistungsfähigkeit von Schülergruppen;
2. in der Lehrerausbildung, in der die Einschätzung von Lehrern durch Schüler thematisiert und angemessene Reaktionen auf diese Einschätzungen erarbeitet werden sollten;
3. von einzelnen Lehrern, die eine Rückmeldung über ihr Verhalten bekommen möchten, die sie aufgrund der Eichung besser als ein eigenes ad-hoc-Verfahren interpretieren und einordnen können.

LEHRER-BEWERTUNGSBOGEN für Schüler

Schultyp: \_\_\_\_\_ Lehrer: \_\_\_\_\_ Geschlecht des Lehrers: \_\_\_\_\_

Klasse: \_\_\_\_\_ Fach: \_\_\_\_\_ Geschlecht des Schülers: \_\_\_\_\_

Welche Zensur hast Du in diesem Fach? \_\_\_\_\_

Wir bitten Dich bzw. Sie hier um eine Beurteilung Deines/Ihres Lehrers. Zwischen zwei gegensätzlichen Aussagen, z. B. zensiert gerecht – zensiert ungerecht sind 7 Kästchen angeordnet. Setze ein Kreuz in eines der Kästchen, wie es Deiner Meinung über den Lehrer entspricht, z. B. ganz links, wenn er oder sie Deiner Meinung nach sehr gerecht zensiert, in das mittlere Kästchen, wenn Du meinst, daß er gerade in der Mitte zwischen gerechtem und ungerechtem Zensieren liegt, und ganz rechts, wenn Du glaubst, daß er sehr ungerecht zensiert. Mit den dazwischenliegenden Kästchen kannst Du noch feinere Abstufungen vornehmen.

Dieser Lehrer/diese Lehrerin

1	zensiert gerecht	<input type="checkbox"/>	zensiert ungerecht					
2	drückt sich verständlich aus	<input type="checkbox"/>	drückt sich unverständlich aus					
3	hat auch außerhalb des Unterrichts Zeit für Schüler	<input type="checkbox"/>	hat außerhalb des Unterrichts keine Zeit für Schüler					
4	läßt Schülermeinungen gelten	<input type="checkbox"/>	läßt nur seine eigene Meinung gelten					
5	kommt immer gut vorbereitet in den Unterricht	<input type="checkbox"/>	kommt nie gut vorbereitet in den Unterricht					

Auf der Rückseite geht es weiter!

6	läßt Schüler über den Unterricht mitbestimmen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	läßt Schüler über den Unterricht nicht mitbestimmen
7	beherrscht sein Fach gut	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	beherrscht sein Fach schlecht
8	zeigt Verständnis für die Probleme der Schüler	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	zeigt kein Verständnis für die Probleme der Schüler
9	ist freundlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ist unfreundlich
10	sein Unterricht ist interessant	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	sein Unterricht ist langweilig
11	ist ausgeglichen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ist launisch
12	behält immer die Ruhe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	regt sich leicht auf
13	sein Unterricht ist systematisch	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	sein Unterricht ist planlos
14	gibt sich Mühe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ist gleichgültig
15	fördert schwächere Schüler	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	gibt schwächere Schüler auf
16	ist beliebt bei den Schülern	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ist unbeliebt bei den Schülern
17	geht auf Anregungen der Schüler ein	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	hält stur an einem vorgefaßten Plan fest
18	verhält sich selbstsicher	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	verhält sich unsicher
19	kann Probleme gut erklären	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	kann Probleme nicht gut erklären
20	führt den Unterricht locker	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	führt den Unterricht straff
21	zensiert milde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	zensiert streng
22	bei ihm kommen alle mit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	bei ihm kommen nur die besten mit
23	man kann ihm vertrauen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	man kann ihm nicht vertrauen
24	ermutigt die Schüler	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	entmutigt die Schüler
25	man lernt viel bei ihm	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	man lernt wenig bei ihm
26	schafft eine zwanglose Atmosphäre	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	schafft eine gespannte Atmosphäre
27	führt den Unterricht sachlich	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	zieht gern eine Show ab
28	hat Sinn für Humor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	hat keinen Sinn für Humor
29	bleibt immer bei der Sache	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	schweift leicht ab
30	nimmt die Schüler ernst	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	nimmt die Schüler nicht ernst
31	sein Unterrichtsziel ist klar zu erkennen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	man weiß nicht, auf was er (sie) hinaus will

32	hält gute Disziplin	<input type="checkbox"/>	hält keine Disziplin							
33	gibt wenig Hausaufgaben auf	<input type="checkbox"/>	gibt viel Hausaufgaben auf							
34	fördert die Selbständigkeit der Schüler	<input type="checkbox"/>	hemmt die Selbständigkeit der Schüler							
35	stellt geringe Anforderungen	<input type="checkbox"/>	stellt hohe Anforderungen							
36	ist mir sympathisch	<input type="checkbox"/>	ist mir unsympathisch							

Vielen Dank fürs Mitmachen!

**Literatur**

Amidon, E. J. und Hough, J. B. (ed.): Interaction Analysis – Theory, Research and Application, Addison-Wesley 1967

Flanders, N. A.: Analyzing Teaching Behavior, Addison-Wesley 1970

Herrmann, Th. und Stapf, A.: Erziehungsstil und Reaktionseinstellung. Bericht an die DFG, Marburg 1968

Herrmann, Th.; Schwitajewski, E. und Ahrens, H. J.: Untersuchungen zum elterlichen Erziehungsstil: Strenge und Unterstützung, Arch. ges. Psychol. 1968, 120, S. 74–105

Herrmann, Th.; Stapf, A. und Krohne, W.: Die Marburger Skala zur Erfassung des elterlichen Erziehungsstils, Diagnostica, 1971, 17, S. 118–33

Hofer, M.: Lehrerverhalten aus der Sicht des Schülers, Unterrichtswissenschaft 1982, S. 240–51

Kahl, Th. N.; Buchmann, M.; Witte, E. H.: Ein Fragebogen zur Schülerwahrnehmung unterrichtlicher Lernsituationen. Zsch. für Entwicklungspsychol. und Pädagog. Psychol., 1977, 9, S. 277–85

Lienert, G. A.: Testaufbau und Testanalyse, Weinheim 1969

Marsh, H. W.; Fleiner, H.; Thomas, C. S.: Validity and usefulness of student evaluations of instructional quality. Journal of Educational Psychology, 1975, S. 833–39

Masendorf, Tücke, Kretschmann, Bartram: Dortmunder Skala zur Lehrerbeurteilung durch Schüler (DSL), Braunschweig 1976. Siehe auch: Zsch. für Entwicklungspsychol. und Pädagog. Psychol. 1977, 9, S. 41–50

Ryans, D. G.: Characteristics of Teachers, Washington 1964

Tausch, R. und A.: Erziehungspsychologie, Göttingen 1977

Stapf, K. H.; Herrmann, Th.; Stapf, A. und Stäcker, K. H.: Psychologie des elterlichen Erziehungsstils, Bern 1972

**Verfasser:**

Dr. Klaus Steltmann, Institut für Erziehungswissenschaft, Universität, Am Hof 3–5, D-5300 Bonn 1