

# Schülerbeurteilungen. Der Berner Beurteilungsbogen im Vergleich mit den Schulnoten der vierten Klasse

Autor(en): **Znoj, Hansjörg**

Objektyp: **Article**

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**

Band (Jahr): **12 (1990)**

Heft 2

PDF erstellt am: **13.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-786262>

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

## **Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

# Schülerbeurteilungen. Der Berner Beurteilungsbogen im Vergleich mit den Schulnoten der vierten Klasse

Hansjörg Znoj

*Dieser Artikel beschäftigt sich mit einem Beurteilungsverfahren, welches in der Stadt Bern seit längerem in Gebrauch ist und als zusätzliche Hilfe dem Lehrer die Entscheidung für einen allfälligen Übertritt eines Schülers/einer Schülerin in die Sekundarschule erleichtern soll. Die Untersuchung gliedert sich in zwei Teile: einerseits wird die Reliabilität des Bogens geprüft, andererseits wird ganz allgemein die Voraussagekraft sowohl der Schulnoten als auch des notenfremen Verfahrens anhand der Leistungen in der Sekundarschule geprüft. Die Ergebnisse sprechen sowohl für zusätzliche Kriterien von Schülerbeurteilungen als auch gegen eine absolute Selektion der Schüler, da die Prognose von Schulleistungen aufgrund der Ergebnisse, die in der vierten Klasse erzielt werden, nicht befriedigen kann.*

## Hintergrund der Untersuchung

Im Rahmen des Aufnahmeverfahrens in die Sekundarschule wird in der Stadt Bern seit längerer Zeit der sogenannte Berner Beurteilungsbogen (Hürsch & Preisig, 1975) versuchsweise auf freiwilliger Basis verwendet. Er setzt sich aus 54 Fragen (Items) zusammen und misst auf neun Dimensionen (Umstellfähigkeit, Mathematisches Denken, Selbständigkeit, Beteiligung, Arbeitsverhalten, Ordentlichkeit, Belastbarkeit, Sprachliches Überlegen und sprachliche Eigenständigkeit)<sup>1</sup> das Verhalten der Schüler. Bei dem Berner Beurteilungsbogen handelt es sich um einen Beobachtungsbogen, der in der Regel vom Lehrer/von der Lehrerin allein ausgefüllt wird. Grosser Wert wurde von den Entwicklern auch auf Aspekte gelegt, die nicht benotet werden können, so Selbständigkeit, Beteiligung und Belastbarkeit. Die genannten Autoren befinden sich damit in Übereinstimmung mit anderen Verfassern von diagnostischen Beobachtungssystemen (Janowski, Fittkau & Rauer, 1981; Besser & Ziegenspeck, 1976).

Der BB (Berner Beurteilungsbogen) hat zum Zweck, dem Lehrer/der Lehrerin eine zusätzliche Hilfe zur Beurteilung eines allfälligen Sekundarschulübertrittes (in der 4. oder in der 5. Primarschulklasse) eines Schülers/einer Schülerin anzubieten. Aufgabe und Ziel der vorliegenden Arbeit war es, den BB im Auftrag der Schuldirektion der Stadt Bern den heutigen Anforderungen anzupassen<sup>2</sup>. In bisherigen Untersuchungen (Romano, 1981; Preisig & Hürsch, 1976; Hürsch, 1977, Hürsch & Aebi, 1976) wurden vor allem die Gründe, die zu einer Entwicklung des BB geführt hatten, aufgeführt.

Als Grundlage der vorliegenden Untersuchung dienten die Daten der Beurteilungsbogen der Jahrgänge 1984, 85, 86, 87 und 88, die auf Lochkarten gespeichert waren. Ausserdem standen die Beurteilungen einiger Schüler aus früheren Jahrgängen zur Verfügung, welche vor allem für die Untersuchung der Schullaufbahnen von grosser Bedeutung waren.

Der Kanton Bern ist zurzeit mit dem Kanton Jura der einzige, der den Übertritt in den anderen Schultyp in der 4. Klasse festlegt, allerdings dürfte dieser Zustand nicht mehr allzulange währen, hat doch das Berner Stimmvolk einem Model 6/3 zugestimmt. Im Kanton Bern lassen sich drei verschiedene Übertrittsmodi feststellen:

- Verfahren, die ganz auf eine Prüfung verzichten, dafür aber zusätzlich zur Leistungsbeurteilung in der vierten Klasse eine Eignungsabklärung mittels zusätzlicher Instrumente (etwa dem BB) vornehmen.
- Verfahren, die «problemlosen» Schülern den prüfungsfreien Übertritt ermöglichen und
- Verfahren, die alle Schüler in die Selektion einbeziehen und Leistungsprüfungen für alle Schüler durchführen.

Zu dem uneinheitlichen Übertrittsverfahren kommt als zusätzlicher Kritikpunkt eine mangelnde Durchlässigkeit. Laufende Schulversuche (Manuel, Bern-West und Twann-Ligerz-Tüscherz) sollen Auskunft über mögliche Reformen geben (Vögeli, 1985).

### **Fragestellung**

Bei der Untersuchung wurde von zwei grundsätzlichen Fragen ausgegangen:

- a) Inwieweit decken sich die Fragen überhaupt mit den postulierten neun Dimensionen des BB? Oder: Misst der Beurteilungsbogen überhaupt das, was er zu messen vorgibt (Konstrukt-Validität)?
- b) Wie genau oder wie gut misst der BB die theoretisch angenommenen Verhaltensmerkmale (Reliabilität)?
- c) Wieviel prädiktorische Validität besitzt der BB? Wieweit deckt sich der Wert des BB (unter Wert des BB ist der Notenwert für Empfehlung im Ergebnisblatt des BB gemeint, dieser Notenwert ist das Ergebnis der Verrechnung der

neun Dimensionen des BB) mit den späteren Noten; wie gut kann der BB zwischen späteren Sekundar- oder Primarschülern trennen (Gültigkeit der Zuweisungsentscheide)?

Vor allem die Frage nach der Gültigkeit ist von entscheidender Bedeutung, werden doch mit einer Selektion für die Zukunft der Schüler Weichen gestellt, die oft später schwer zu korrigieren sind. So wird zwar im Kt. Luzern seit 1977 ein Übertrittsverfahren angewandt, welches aus den drei Elementen Prüfung, Schülerbeurteilung und den Erfahrungsnoten besteht. Diese drei Elemente ergeben eine «Übertrittsnote», die Selektionsnormen sind fixiert. Trotz der hohen prognostischen Validität dieses Verfahrens kommt Casparis (1984) zum Schluss, dass die Selektion als solche mit immanenten Schwierigkeiten verbunden ist. So steigt beispielsweise das Anforderungsniveau bei den Prüfungen (Leistungseskalation), damit gewährleistet werden kann, dass die traditionelle Quotenverteilung erhalten bleiben kann. Verlässliche Ausleseverfahren sind nur unter bestimmten Bedingungen möglich: die Ausbildungsziele müssen bestimmt werden, es braucht eindeutige Erfolgskriterien und setzt eine relative Konstanz der Persönlichkeitsvariablen sowie die Messbarkeit dieser Persönlichkeitsvariablen voraus. «Das schweizerische Schulsystem ... gründet auf der (vorwissenschaftlichen) Annahme, dass die Ausleseverfahren mit hoher Verlässlichkeit die Chance, Ausbildungsziele zu erreichen, vorhersagen.» (Trier, 1974, S. 1001). Im allgemeinen konnte nachgewiesen werden, dass die Prognosen von Schulleistungen mit Noten schwach sind. Trotzdem haben laut Trier (1974) Noten höhere prognostische Validität als Prüfungen. Eine Hauptschwierigkeit der Notengebung dürfte deren Funktion sein. Flammer, Perrig und Rüegg (1983) unterscheiden fünf verschiedene Funktionen:

- Berechtigung (für Schule, Titel, Studium, Übertritt etc.)
- Selektion
- Orientierung
- Motivierung
- Beurteilung

Der Lehrer/die Lehrerin hat also mindestens eine Doppelaufgabe zu bewältigen, einerseits hat er/sie den Schüler/die Schülerin zu fördern, andererseits hat er/sie eine Diagnose zu leisten, also eine Bewertung durchzuführen. Nach Flammer et al. (1983) ist vom heutigen Schulsystem einzig die summative Evaluation zuhanden des Promotionsentscheides die einzige unabdingbare Funktion der Schulnote. Die Noten dienen demnach als administratives Entscheidungsmittel. Gerade diese Funktion wird jedoch gegenwärtig stark diskutiert und ist nicht unumstritten, zeigen doch Untersuchungen im Kanton Bern (Brunner, 1989) und im Kanton Baselland (Halbherr, 1981) dass der durchschnittliche Real- (oder Primar-) Schüler in Leistungstests besser abschneidet als ein schlechter Sekundarschüler.

## Methoden

Die erste Frage, inwieweit sich die Fragen des BB<sup>3</sup> mit den postulierten Dimensionen decken, wurde faktorenanalytisch geprüft. Gerechnet wurde mit

dem SPSS-Programm auf dem zentralen Rechner der BEDAG (Rechenzentrum Bern). Als Grundlage dienten die auf Lochkarten gespeicherten Daten der Jahre 1984–88. Es wurden die Beurteilungen von 617 Schülern verrechnet.

Der Grad der Zuverlässigkeit (Reliabilität) womit der BB ausgefüllt wird, wurde mit dem Test «Reliability» ausgerechnet. Das verwendete CRONBACH ALPHA ist eine varianzanalytische Reliabilitätsschätzung, wobei Reliabilität als Anteil der «wahren Varianz» an der Gesamtvarianz definiert ist.

Die Zusammenhänge zwischen BB-Werten und Schulnoten oder Noten mit Noten wurden mit einem nichtparametrischen (verteilungsfreien) Verfahren gerechnet. Verwendet wurde der Spearman'sche Rangkorrelationskoeffizient. Die Stichprobengrösse betrug 78 Schüler. Die Auswahl der Schüler war zufällig bis auf die Jahre 1979, 1982 und 1983, da in diesen Jahren nur wenige Daten vorhanden waren. Aus den Jahren 1984, 1985 und 1986 wurden je 20 Schüler aus verschiedenen Schulen (und mit verschiedenen Lehrern/Lehrerinnen) zufällig herausgegriffen.

Im Zentralregister der Schuldirektion wurden die Schulhäuser eruiert, welche die betroffenen Schüler und Schülerinnen zu der gefragten Zeit besucht hatten, und über die Schulen wurden dann folgende Noten erhoben: Deutsch, Rechnen und Heimatkunde für die 4. Klasse sowie Deutsch, Französisch, Mathematik, Geschichte und Geographie für die folgenden Klassen. Nach der Erhebung der Noten erhielt jeder Schüler einen Zahlencode, die Daten wurden so anonymisiert.

Der letzte Fragetypus, z. B., ob die Promotionsnoten der 4. Klasse oder die Werte des BB besser zwischen Primar- oder Sekundarschule zu trennen vermögen, wurde mittels t-Tests gerechnet. Das Signifikanzniveau wurde auf 5% zweiseitig festgelegt. Daneben wurden die Zuweisungsentscheide für die 4. und 5. Klasse getrennt aufgelistet.

Als fehlend wurden folgende Angaben betrachtet: Wechsel in eine Privatschule<sup>4</sup>, Wegzug aus dem Kanton, nicht erreichte Schulstufe. Bei allfälligen Repetitionen wurde ausschliesslich das erste Zeugnis berücksichtigt. Ein Wechsel von der Primarschule in die Sekundarschule später als in der 5. Klasse wurde gleichgesetzt mit Sekundarschule, eine Zurückstufung mit Primarschule (nur ein Fall). Als Zufallsfaktor muss die Tatsache gelten, dass in der vorliegenden Stichprobe überdurchschnittlich viele 5.-Klasse-Beurteilungen vorliegen, die Erfahrung der jährlichen Auswertung hätte eine gleichmässige Verteilung erwarten lassen.

## Ergebnisse

### 1) Ergebnisse aus der Faktorenanalyse

Die nach der Varimax-Methode (Kaiser-Normalisierung) rotierten Items des Beurteilungsbogens (n = 671) ergaben schliesslich sechs Faktoren mit einem Eigenwert grösser als eins. Die sechs Faktoren erklären zusammen 71,3% der Gesamtvarianz, das heisst, dass fast  $\frac{3}{4}$  der Streuung des Ratings der 54 Fragen

mit sechs unabhängigen Faktoren erklärt werden kann. Die sechs Faktoren können wie folgt beschrieben werden:

Faktor 1 46% der Gesamtvarianz

Dieser Faktor deckt etwa folgenden Merkmalsbereich ab: Umstellfähigkeit, Abstraktionsvermögen, logisches Denken.

Faktor 2 8,3% der Gesamtvarianz

Faktor 2 beinhaltet die Merkmale Selbständigkeit, Beteiligung, Kreativität und Komplexität.

Faktor 3 7,2% der Gesamtvarianz

Der Faktor 3 lässt sich wie folgt umschreiben: Ordnungssinn, Konzentration.

Faktor 4 4,8% der Gesamtvarianz

Der Faktor 4 deckt den Bereich Belastbarkeit und Stressresistenz ab.

Faktor 5 3,0% der Gesamtvarianz

Der Faktor 5 lässt sich als Sprachliches Überlegen und Phantasie interpretieren.

Faktor 6 2,0% der Gesamtvarianz

Der Faktor 6 besteht nur aus einer Frage (54): «Er hat Sinn für Humor» (lädt auch auf Faktor 2).

Der Faktor 1 (Umstellfähigkeit, Abstraktionsvermögen, logisches Denken) erklärt den Hauptanteil der Gesamtvarianz. Alle anderen Fragen tragen nicht so viel bei zur Differenzierung der einzelnen Schüler; immerhin erhöht sich der Anteil an erklärter Varianz auf 71,3%, wenn zusätzlich die anderen fünf Faktoren dazugenommen werden.

Die Diskrepanz zu den postulierten neun Dimensionen des BB soll nicht weiter erstaunen: Die Entwickler des BB sind von theoretisch fundierten Dimensionen ausgegangen, d.h. es lag gar nicht im Sinne der Autoren des BB, von mathematisch unabhängigen Faktoren auszugehen. Deshalb wurde auch eine Faktorenanalyse mit neun (nicht orthogonalen) Faktoren gerechnet. Das Ergebnis dieser Analyse konnte jedoch nicht sinnvoll interpretiert werden, d.h. die einzelnen Fragen (Items) laden auf Faktoren, die inhaltlich widersprüchlich zu den Items sind. Das verwundert insofern wenig, als Hürsch und Mitarbeiter auch gar nicht im Sinne hatten, den BB nach den einzelnen Dimensionen auszuwerten: Bei der Auswertung des BB wird jede Frage einzeln mit einem Faktorenwert multipliziert<sup>5</sup>. Bereits Hürsch (1977) hatte mit den 1976 erhobenen Daten eine Faktorenanalyse gerechnet und war auf sechs orthogonale Faktoren gekommen, die sich inhaltlich weitgehend mit den oben angeführten zur Deckung bringen lassen. Die Population von 1976 lässt sich demnach vergleichen mit der Stichprobe von 1984–1988.

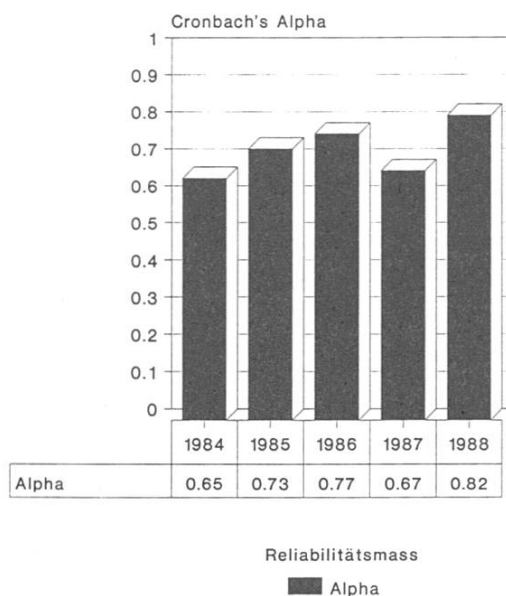
## 2) Die Reliabilität des Berner Beurteilungsbogens

Ausgehend von einem  $n = 617$  wurde eine varianzanalytische Reliabilitätschätzung nach Cronbach gerechnet. An Stelle des «wahren Wertes» tritt der universe score  $M$ , der definiert ist als Mittelwert aller beobachteten Werte über

alle möglichen Bedingungen eines inhaltlich umschriebenen «Universums» (hier die Gesamtstichprobe). Die Frage ist nun, wie gut von einer bestimmten Bedingungskondition aus auf diesen universe score generalisiert werden kann. Das Mass ist der Koeffizient (Alpha).

Ergebnisse:

Tabelle 1  
Reliabilität



n total = 671

(1984 n = 164; 1985 n = 115; 1986 n = 132; 1987 n = 140; 1988 n = 120)

Die Gesamt-Reliabilität beträgt ALPHA = .73

Alpha ändert sich, falls einzelne Items aus der Berechnung herausgenommen werden, nicht wesentlich (Alpha zwischen .75 und .71); das heisst: es existiert kein Item, das sich nicht reliabel verhält oder dessen Einschätzung überdurchschnittlich viel Mühe bereitet.

### 3) Zusammenhang des BB-Wertes mit Schulnoten

#### 3.1) Zusammenhang des BB-Wertes mit 4.-Klasse-Noten

In der Gesamtstichprobe (n = 78) finden sich folgende Zusammenhänge (Spearman'scher Rangkorrelationskoeffizient) des BB-Wertes mit den erzielten Noten der 4. Klasse:

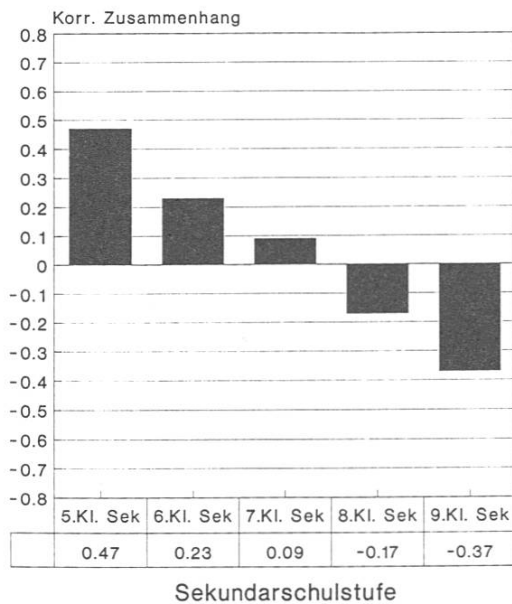
BB mit Deutsch:	.46**	
BB mit Rechnen:	.36**	
BB mit Heimatkunde:	.45**	
BB mit Prom.-schnitt:	.47**	6

Der Zusammenhang zwischen der Beurteilung des Lehrers durch den BB und der Noten der 4. Klasse (durch denselben Lehrer) ist statistisch signifikant. Stärker interessiert jedoch der Zusammenhang des BB-Wertes mit den Noten, die die jeweiligen Schüler in den folgenden Schuljahren erzielten und damit auch die Frage, ob nun der BB oder die Schulnoten der 4. Klasse den weiteren Schulerfolg besser vorhersagen, umso mehr, als der Zusammenhang des BB mit dem Durchschnitt der Promotionsnoten mit .47 auch grösser hätte ausfallen können.

Hier interessiert vor allem der Zusammenhang der Promotionsnoten der 4. Klasse und der Beurteilungen des BB mit den später erzielten Noten in der Sekundarschule (ab der 5. Klasse), umso mehr, als der BB eigens für diesen Zweck konstruiert wurde.

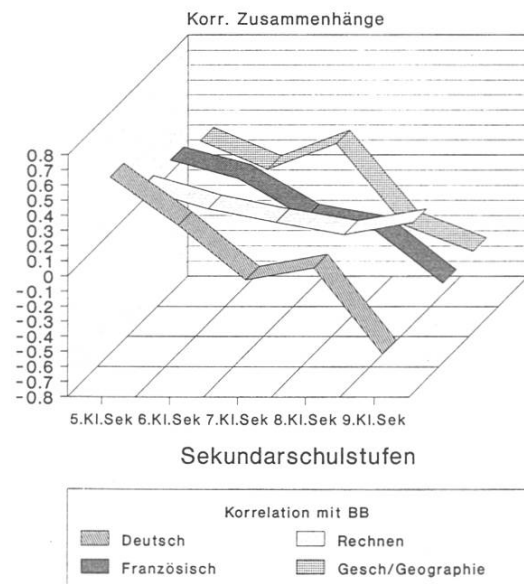
### 3.2) Zusammenstellung des korrelativen Zusammenhangs von BB und den Sekundarschulnoten

Tabelle 2.1  
BB -Werte mit Sek-Durchschnitt



5.Kl.n=25; 6.n=31; 7.n=20; 8.n=12; 9.n=7

Tabelle 2.2  
BB mit Sek.-Noten: D,R,F,G/G

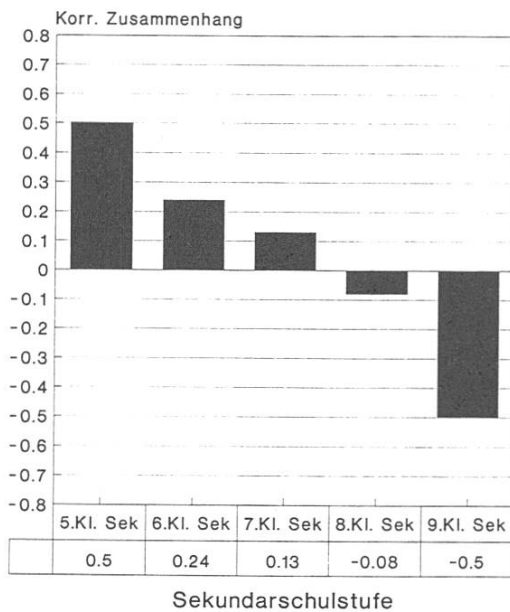


5.Kl.n=26; 6.n=32; 7.n=21; 8.n=13; 9.n=8



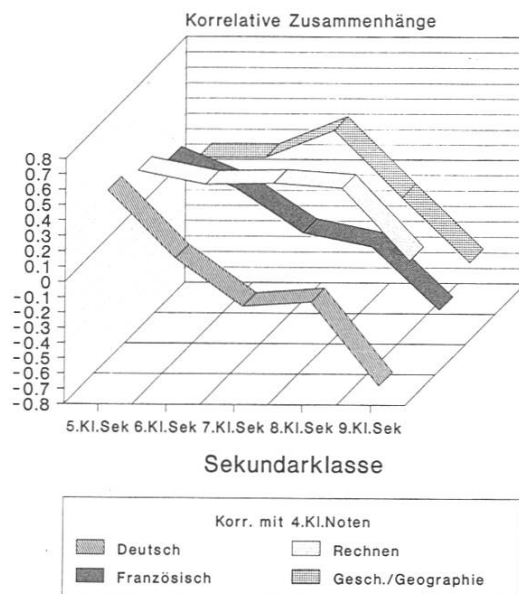
3.3) Zusammenstellung des korrelativen Zusammenhangs zwischen den Noten der 4. Primarklasse und den in den folgenden Sekundarschulstufen erzielten Noten <sup>7</sup>.

Tabelle 3.1  
Prom.-Noten mit Sek-Durchschnitt



5.Kl.n=25; 6.n=31; 7.n=20; 8.n=12; 9.n=7

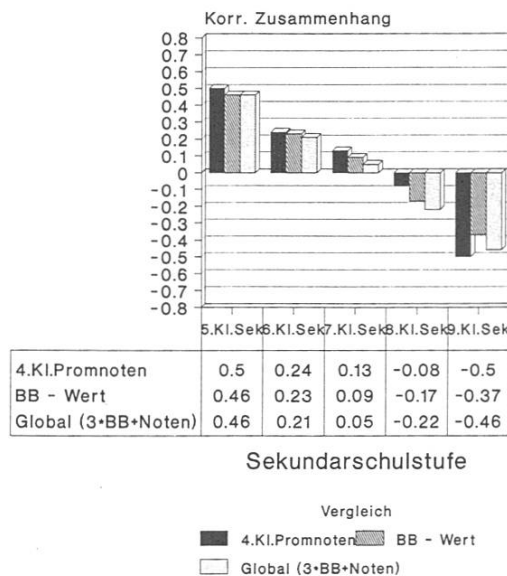
Tabelle 3.2  
Prom.-Noten mit Sek.-Noten: D,R,F,G



5.Kl.n=25; 6.n=31; 7.n=20; 8.n=12; 9.n=7

3.4) Vergleich von Beurteilungsbogen, Promotionsnoten und dem Globalmass, welches sich zu gleichen Teilen aus BB-Wert und Durchschnittsnoten der 4. Klasse zusammensetzt.

Tabelle 4  
Vergleich von BB, Prom & Glob



5.Kl.n=25; 6.n=31; 7.n=20; 8.n=12; 9.n=7

Vergleicht man die Ergebnisse, so sprechen die Resultate weder für die Schulnoten noch für den Beurteilungsbogen, wenn die Voraussage längerfristig gemacht werden soll. Unmittelbar in der 5. Sekundarschulklasse wirken sich die Zusammenhänge mit den 4.-Klasse-Promotionsnoten etwas stärker aus. Als einzelne Note ist einzig die Rechennote relativ stabil bis zur 8. Klasse, dann wird auch hier der Zusammenhang negativ. Zum Teil werden die korrelativen Zusammenhänge in der 9. Sekundarklasse verglichen mit den 4.-Klasse-Noten sogar negativ (Deutsch,  $-.73$ ). Dies ist ebenfalls für den Beurteilungsbogen der Fall (Deutsch,  $-.57$ ). Die Prognose wird auch nicht besser, wenn beide Verfahren, also Promotionsnoten und BB-Wert, kombiniert angewendet werden, wie aus Tabelle 4 deutlich hervorgeht.

#### 4) Zuweisungsentscheide

##### 4.1) Zuweisungsentscheide unter Berücksichtigung von BB-Wert und 4.-Klasse-Promotionsnoten.

###### Ergebnisse:

Sowohl der Beurteilungsbogen als auch alle Schulnoten (Deutsch, Rechnen, Heimatkunde) trennen signifikant (t-Test, Signifikanzniveau zweiseitig 5%)<sup>8</sup> zwischen späteren Primar- und Sekundarschülern. Berücksichtigt man nur die 4.-Klasse-Burteilungen (was deswegen Sinn ergibt, weil nur insgesamt drei Schüler aus der vorliegenden Stichprobe aus der 5. Klasse in die Sekundarschule wechselten), so trennt der BB tendenziell zwar besser als die einzelnen Noten, nicht aber als der Durchschnitt der Promotionsnoten. Der reale Zuordnungsentscheid wird davon aber nicht betroffen.

Tabelle 5.1  
Zuweisungsentscheid 4. Klasse

<u>Kriterium</u>	<u>Prim.</u> (n=20)	<u>Sek.</u> (n=30)
BB - Wert	3.5	4.3
Promotionsschnitt	4.5	5.2
Deutsch	4.5	5.1
Rechnen	4.3	5.0
Heimatkunde	4.7	5.3

Tabelle 5.2  
Zuweisungsentscheid 5. Klasse

<u>Kriterium</u>	<u>Prim.</u> (n=22)	<u>Sek.</u> (n=3)
BB - Wert	4.2	4.4
Promotionsschnitt	4.3	4.6
Deutsch	4.3	4.3
Rechnen	4.3	4.5
Heimatkunde	4.5	5.0

Man darf nach diesen Resultaten davon ausgehen, dass sowohl die Promotionsnoten als auch der Beurteilungsbogen zwischen beiden Gruppen differenzieren, dem Lehrer eine genügend gute Grundlage für den Zuweisungsentscheid liefern. Dabei sollte jedoch nicht unberücksichtigt bleiben, dass das Kriterium nur Noten beinhaltet, insofern also dem «notenfreien» Instrument BB nicht ganz gerecht zu werden vermag. Dies lässt sich allerdings rechtferti-

gen, da ja auch die Selektion praktisch auf den Schulnoten basiert. Die 5.-Klasse-Burteilungen mit dem BB ergeben jedoch, gemessen an den realen Zuweisungsentscheiden, zu hohe Werte.

#### *4.2) Bewertung der Schulnoten hinsichtlich des Schultyps und der Schulstufe*

Die Frage, ob in der neunten Klasse grundsätzlich anders (besser) als in den übrigen Schulklassen beurteilt wird, lässt sich aufgrund der Ergebnisse von t-Tests nicht eindeutig beantworten. Hingegen werden offensichtlich in den beiden Schultypen Leistungen unterschiedlich bewertet: In der Primarschule steigen die Notenwerte über alle Schüler in allen Fächern in allen folgenden Klassen gegenüber den 4.-Klasse-Noten statistisch signifikant an (abhängiger t-Test, 5%-Niveau). Dies ist bei den Sekundarschülern nicht der Fall, dort ist das Noten-Niveau etwa gleich, tendenziell eher geringer bei einem leichten Aufwärtstrend in der neunten Klasse, der sich allerdings statistisch nicht belegen lässt. Bei beiden Schultypen pendelt sich die Durchschnittsnote über alle Schüler bei fünf ein.

### **Diskussion**

Zum jetzigen Zeitpunkt wird der «Berner Beurteilungsbogen» nur von wenigen Lehrern und Lehrerinnen benutzt. Das hat verschiedene Gründe. Einer davon ist der Umfang des Bogens (der allerdings verglichen mit ähnlich konzipierten Instrumenten, z.B. von Janowski et. al., 1981 relativ bescheiden ist), ein anderer Grund besteht sicher darin, dass es für viele Lehrer und Lehrerinnen nicht einsichtig ist, weshalb sie zur differenzierten Beurteilung eines Schülers oder einer Schülerin so viele Fragen beantworten sollen. Die Ergebnisse aus der Faktorenanalyse zeigen, dass sich die 54 Fragen zu sechs, beziehungsweise fünf orthogonalen oder unabhängigen Faktoren subsumieren lassen: allerdings hat dies keinen Einfluss auf eine allfällig andere Gestaltung des BB, weil die Entwickler für die Auswertung jedes einzelne Item mit einem festen Faktorenwert (nicht unabhängige Faktoren) multiplizieren. Der Zuweisungsentscheid fällt beim BB ebenso eindeutig aus wie bei den Schulnoten der vierten Klasse. Es stellt sich also die Frage, was den zusätzlichen Aufwand, der zum Ausfüllen des BB benötigt wird, rechtfertigt. Da ist zum einen das herrschende Unbehagen, einen Schüler/eine Schülerin nur nach Leistungskriterien zu beurteilen. Darüber hinaus hat der BB einen ebenso «hohen» Voraussagewert über einen längeren Zeitraum wie die Promotionsnoten (diese Tatsache kann jedoch nicht über den schwachen Prognosewert des Beurteilungsbogens und insbesondere der Schulnoten hinwegtrösten...). Die Ergebnisse der Reliabilitätsprüfung bescheinigen dem BB für ein Selektionsverfahren sehr hohe und stabile Messgenauigkeit; nach den Ergebnissen des Reliabilitätstests ergeben sich auch keine Fragen (Items), die unterschiedlich genau, verglichen zu den übrigen Fragen, messen. Zu der «Retest-Reliabilität» stellt allerdings Flammer (1975) fest, dass diese mit zunehmendem Abstand der Notennahme (dies gilt offensichtlich

auch für notenfreie Beurteilungsverfahren) sinkt, weil sich Schüler und Schülerinnen tatsächlich verändern. Korrelationen vermengen die Aspekte der persönlichen Entwicklung und der Zuverlässigkeit solcher Erhebungen. Dies gilt selbstverständlich auch für die vorliegende Untersuchung. Der Vorteil des BB gegenüber den Noten dürfte ein anderer sein: die Funktion des BB ist keine erzieherische, sie soll einzig dem Lehrer/der Lehrerin beim Entscheid Primar- oder Sekundarschule helfen. Die Trennung zwischen Beurteilung und Unterricht erfolgt viel deutlicher. So wie die Note momentan gehandhabt wird, sind zuviele Funktionen in ihr vereinigt. Unter Umständen können auch die Schüler selbst zur Beantwortung der Fragen beigezogen werden, wie es laut mündlicher Mitteilung vereinzelt bereits gemacht wird. Mit einem Instrument wie dem BB umgeht man auch den Mythos der «Einheitlichkeit und Eindeutigkeit» der Benotung. Flammer et al. (1983) konnten zeigen, dass die Notenverteilung einer Klasse je nach Lehrer leicht eine halbe bis ganze Note ausmachen kann. Diese Unterschiede im Skalenverständnis der Lehrer (Flammer und Wiegand, 1973) könnten durch ein standardisiertes Training eines Beurteilungsverfahrens behoben werden. Eine viel grundsätzlichere Kritik betrifft freilich die Schulstruktur selbst. Die Selektion im vierten Schuljahr wird allgemein als verfrüht betrachtet, vor allem dann, wenn dazu eine mangelnde Durchlässigkeit kommt. Im Schulversuch Spiegel (Gde. Köniz, Kt. Bern) waren 50% der Schüler entweder dem oberen oder unteren Niveau in den Fächern Deutsch, Französisch und Mathematik zugeteilt. «In der Regelschule müsste somit die Hälfte der Schüler in einem der drei Fächer entweder über- oder unterfordert werden» (Brunner, 1989; dazu auch Ramseier, 1982). Dieser Umstand wird auch nicht durch das zukünftige Strukturmodell 6/3 behoben, das zwar die aktuelle Situation mit Sicherheit entschärfen wird, aber letztlich problematisch bleibt. Die Ergebnisse dieser Untersuchung sprechen jedenfalls einem «durchlässigen» Erziehungssystem das Wort, da sich Schulleistungen offensichtlich nicht genügend stark messen und prognostizieren lassen.

### Anmerkungen

- <sup>1</sup> Beispiele für solche Items sind (Rating: meistens, öfters, gelegentlich, seltener, kaum): «Wird eine bekannte Aufgabe auf eine neue Art gestellt, stutzt der Schüler und hat Verständnisschwierigkeiten» oder «Der Schüler erbringt unter Zeitdruck schlechtere Leistungen als gewöhnlich» oder «Das Verhalten des Schülers ist vor oder während einer Probe (Prüfung) besonders ungeschickt (motorische und verbale Fehlleistungen)» oder «Es gelingt dem Schüler leicht, sich von eingeübten Denkweisen zu lösen» oder «Er sagt seine Meinung, auch wenn er riskiert, dass ihn die Klasse auslacht». (Eine geschlechtsneutrale Schreibweise wird für die Zukunft angestrebt.)
- <sup>2</sup> Häufigste Kritikpunkte der Lehrerschaft betrafen einerseits einzelne Fragen aus dem Beurteilungsbogen (zu vage, zu unscharf, zu komplex...); andererseits lautete eine globale Kritik, der Bogen bedeute für den Lehrer/die Lehrerin eine zu hohe zeitliche Belastung, er sei zu lang.
- <sup>3</sup> Der Fragebogen setzt sich aus 54 Items zusammen, die fünfstufig skaliert sind (meistens, öfters, gelegentlich, seltener und kaum). Der Fragebogen ist erhältlich in der städtischen Schuldirektion Bern.

- <sup>4</sup> Ich glaube auf eine detaillierte Notenangabe aus Privatschulen aus folgenden Gründen verzichten zu können: einerseits scheinen mir vor allem soziale Motive für einen Wechsel in eine Privatschule massgebend zu sein, andererseits handelt es sich durchwegs um Schüler, die sowohl von den Schulnoten als auch von den Werten des BB als nicht geeignet für die Sekundarschule betrachtet worden sind.
- <sup>5</sup> Dieser Befund spricht gegen eine Kürzung des BB, da man nicht einzelne Fragen aus dem Gesamtgefüge entfernen kann, ohne den Bogen als solches völlig neu zu gestalten.
- <sup>6</sup> zwei \*\* bedeuten: statistisch signifikant auf 1%-Niveau  
ein \* bedeutet: statistisch signifikant auf 5%-Niveau
- <sup>7</sup> Der Vergleich findet ausschliesslich zwischen den Primarschulnoten der 4. Klasse und den Sekundarschulnoten der 5. und fortfolgenden Klassen statt, da ja die erzielten Notenwerte nur für diesen Fall von Interesse sind.
- <sup>8</sup> Die t-Test-Werte sind auf dem 1%-Niveau signifikant. Im einzelnen betragen sie:  
BBt = 4.51\*\*; Deutsch t = 3.39\*\*; Rechnen t = 3.81\*\*; Heimatkunde t = 3.99\*\*  
und für den Promotionsdurchschnitt t = 4.56\*\*.

### Literaturangaben

- Besser, H. & Ziegenspeck, J. (1976). Verfahren zur Schülerbeobachtung. In: schulmanagement 1, 64–66.
- Brunner, J. (1989). Plädoyer für eine integrierte Oberstufe. schweizer schule, 4,3–14.
- Casparis, C. (1984). Eine Evaluation des Luzerner Übertrittsverfahrens. Bildungsforschung und Bildungspraxis, 1,17–29.
- Flammer, A. (1971). Zur Definition der Notenskala. Schw. Zeitschrift f. Psychologie, 30, 204–218.
- Flammer, A. (1975). Individuelle Unterschiede im Lernen. Basel: Beltz.
- Flammer, A. & Wiegand, E. (1973). Schulnotenreliabilität und ihre Minderungsfaktoren. Schweiz. Zeitschrift f. Psychologie, 32, 46–55.
- Flammer, A.; Perrig-Chiello, P.; Rüegg, T. (1983). Zeugnisnoten vor dem Übertritt in die Sekundarstufe. Freiburg: Universitätsverlag.
- Halbherr, P. (1981). Die Selektion in der Oberstufe. Basellandschaftliche Schulnachrichten, 18–26.
- Hürsch, L. (1977). Fopsy-Press. Forschungsabteilung für Begabungs- und Bildungsfragen. Universität Bern.
- Hürsch, L. & Aebi, H.J. (1976). Analysen von Aufnahmeverfahren. Die Aussagekraft von Primarschulnoten und weiteren Aufnahmekriterien für den Besuch der Sekundarschule. Bericht 2: Stadt Bern II, Bern.
- Hürsch, L. & Preisig, E. (1975). Entwicklung und Evaluation eines Schülerbeobachtungsbogens zum Zweck der Selektion von Schülern. Unveröffentlicht, Universität Bern.
- Janowski, A.; Fittkau, B. & Rauer, W. (1981). Beurteilungshilfen für Lehrer. Westermann Test, Braunschweig.
- Preisig, E. & Hürsch, L. (1976). Zuverlässigkeit und Gültigkeit des Berner Beurteilungsbogens. Unveröffentlicht, Universität Bern.
- Ramseier, E. (1982). Schulversuch Manuel. Bern: Haupt.
- Romano, R. (1981). Die Vorhersagbarkeit des Schulerfolgs. Untersuchungen über die prädiktorische Validität des Berner Beurteilungsbogens. Unveröffentlichte Lizentiatsarbeit, Bern.
- Trier, U. (1974). Die Problematik von Entscheidungsgrundlagen beim Übertritt von der Primarschule in die weiterführende Schule. Schweizer Schule, 22, 999–1004.
- Vögeli, U. (1985). Von der Primarschule in die Sekundarstufe I in der Schweiz. Projekt SIPRI, EDK, Bern.

## L'évaluation des élèves. Le «Questionnaire d'évaluation bernois» comparé aux notes de 4<sup>e</sup>.

### *Résumé*

Cet article porte sur un instrument d'évaluation utilisé depuis quelque temps à Berne comme moyen auxiliaire et destiné aux maîtres pour faciliter la décision quant au passage des élèves dans l'enseignement secondaire. La recherche se divise en deux parties: nous avons évalué d'une part la fidélité du questionnaire, d'autre part, de façon très générale, la validité pronostique aussi bien des notes primaires que de la procédure d'évaluation sans note par rapport à la réussite à l'école secondaire. Les résultats parlent aussi bien en faveur de critères supplémentaires pour l'évaluation des élèves que contre une sélection absolue de ceux-ci. En effet, la prévision de la réussite scolaire ultérieure sur la base des résultats de 4<sup>e</sup> n'est pas satisfaisante.

## Judgements of pupils. The «Berner Beurteilungsbogen» in comparison with the marks of the fourth class in primary school.

### *Summary*

This study is about a rating procedure, called «Berner Beurteilungsbogen», which is in use in the city of Bern. It is meant to help the teacher to decide whether a pupil is to be proposed for secondary school or not. The study is divided in two parts: on the one hand the reliability of the procedure is tested, on the other hand the predictive accuracy for the marks in primary school as for the rating procedure is correlated by the marks in secondary school. The results speak for additional criteria with regard to judgements of pupils and against an absolute selection as the predictive accuracy based on the achievement in the fourth class is rather low in the following classes in secondary school.