

Originalarbeiten

Instrumentalitätstheoretische Analysen von schulischen Benotungsprozessen – Eine empirische Pilot-Studie

Günter Krampen, Petra Lehmann und Margarete Haag

Zusammenfassung: In der vorliegenden Arbeit werden Handlungspräferenzen für die Vergabe bestimmter Zeugnisnoten in problematischen Benotungssituationen auf der Basis eines einfachen instrumentalitätstheoretischen Modells, das Bezug nimmt auf (a) Handlungsziele und deren Bewertungen durch das Handlungssubjekt, (b) Handlungsalternativen (Notenstufen) und ihr subjektiv erwarteter Beitrag zur Erreichung der Handlungsziele, analysiert. 51 berufserfahrene Lehrer und 87 Lehramtskandidaten bearbeiteten einen Fragebogen, in dem u.a. nach Erziehungszielbewertungen, dem Benotungsverhalten in bestimmten Situationen und der Wahrnehmung dieser Situationen gefragt wurde. Die durchgeführten Analysen bestätigen die Tauglichkeit des Ansatzes zur Analyse von Benotungsverhalten, machen jedoch auch deutlich, daß die befragten Lehrpersonen Noten nur sehr bedingt einen Beitrag zur Erreichung hoch bewerteter pädagogischer Zielsetzungen zuschreiben. Die Ergebnisse zu den erhobenen Variablen werden im einzelnen dargestellt; auf die Bedeutung der Analyse von Benotungsverhalten wird eingegangen.

Instrumentality-theoretical analysis of grading behavior in school – An empirical pilot-study

Summary: In the present study, an action-theoretical model is proposed which treats action preferences for grading behavior in problematical grading situations as a function of individual goalvaluations and of perceived instrumentalities of action alternatives (marks) for goal attainment. 51 school teachers and 87 student teachers filled out a questionnaire, which included educational goals, grading behavior in special situations, perception of these situations etc. Data analysis substantiate the usefulness of the model for studying grading behavior, but shows too, that teacher attribute only limited contributions to marks for reaching high valued educational goals. Results for all collected variables are presented in detail; the relevance of analysis of grading behavior is pointed out.

1. Problemstellung

Das Benotungsverhalten von Lehrern in der Schule wurde bislang vor allem an Hand der sozialwissenschaftlichen Forschungskriterien Objektivität, Reliabilität und Validität untersucht. Zu den beiden erst genannten Kriterien liegen eine relativ lange Forschungstradition (vgl. etwa Eelss 1930, Finalyson 1951) und reichhaltige Befunde vor, die darauf hinweisen, daß Lehrerurteile

über die Schulleistungen von Schülern nur eine geringe Objektivität, Vergleichbarkeit und Reliabilität besitzen (vgl. hierzu zusammenfassend Ingenkamp 1971, Stephan 1978). Ähnliches gilt für das Kriterium der Validität (vgl. hierzu im Überblick etwa Zielinski 1974), wobei allerdings zu berücksichtigen ist, daß die durch den Lehrer und das Schulsystem intendierten Funktionen von schulischen Zensuren bei der Auswahl der Validierungskriterien in empirischen Untersuchungen meist vernachlässigt wurden.

Bei diesen Untersuchungen zum Benotungsverhalten von Lehrern geht es nun im Grunde um die *Bewertung* der Genauigkeit, Zuverlässigkeit, Vergleichbarkeit und Gültigkeit des Lehrerurteils, nicht jedoch um die *Analyse* des Zensierungsverhaltens selbst. Solche analytischen Studien, die vor allem auf das Aufdecken von Fehlerquellen im Lehrerurteil ausgerichtet sind, beschäftigen sich bislang schwerpunktmäßig mit den Wahrnehmungsfehlern des Lehrers (vgl. etwa Knoche 1969, Heller 1974, im Überblick Höger 1978), speziell mit den sozialpsychologischen Variablen Vorurteil (vgl. etwa Hadley 1954, Steinkamp 1968, Ferdinand 1971) und Stereotyp (vgl. etwa Schäfer 1975), die unter bestimmten Randbedingungen eine besondere Bedeutung erlangen (etwa bei unterschiedlicher Schichtzugehörigkeit der Schüler, siehe Hanke, Lohmöller, Mandl 1970, oder bei Fachleistungsdifferenzierungen, siehe dazu auch Knapp 1978). Neuerdings rücken die kognitiven Bedingungen der Benotung bei Lehrern in den Blickpunkt des Forschungsinteresses. Hier sind vor allem Arbeiten zu den impliziten Theorien des Lehrers über die Schülerpersönlichkeit (vgl. Kleiter 1973, Hofer 1974, Masendorf, Tscherner, Tücke 1974; Mandl Lohmöller, Hanke 1975), zur Bezugsnormorientierung (vgl. etwa Rheinberg 1976, Rheinberg, Schmalt, Wasser 1978) und zum Attribuierungsverhalten von Lehrern zu nennen (vgl. etwa Meyer, Butzkamm 1975; Herrig, Krampen 1979).

Es stellt sich nun die Frage, ob das Benotungsverhalten von Lehrern nur von den beobachteten Schülerleistungen, Wahrnehmungsfehlern unterschiedlichster Art, den impliziten Persönlichkeitstheorien des Lehrers, seinen Kausalattributionen und seiner Bezugsnormorientierung abhängt, oder ob zusätzlich weitere Variablenkomplexe dieses Verhalten mitbestimmen. Betrachtet man die Kataloge von Funktionen, die schulischen Zensuren *zugeschrieben* werden, also bislang nicht Gegenstand empirischer Arbeiten waren¹, so zeigt sich, daß neben Zuweisungen wie Rückmeldefunktion für Schüler, Eltern und Lehrer, Sozialisierungsfunktionen, Selektions- und Zuteilungsfunktionen stets auch im eigentlichen Sinne „pädagogische“ Funktionszuweisungen getroffen werden. Noten sollen danach eine „Anreizfunktion“ für den Schüler haben (Zielinski 1974, p. 881), sie haben die Funktionen der „extrinsischen Motivierung“ und der Disziplinierung von Schülern (Ingenkamp 1975, p. 13; Kleber 1976, p. 28ff.), sie dienen allgemein der „pädagogischen Intervention“ (Stephan 1978, p. 14). Inwieweit werden nun diese pädagogischen Funktionen von Noten bei der Notengebung beachtet? Sind pädagogische Diagnostik und Zensuren tatsächlich zu „einer Art Selbstzweck“ (Meister 1976, p. 225) geworden, steht die *Technik* des Beno-

tungsverhaltens (also Objektivität, Reliabilität, Validität, Normierungen, soziale Bezugsnormbeachtung und Mittelwertbildungen etc.) im Vordergrund und werden dadurch eventuell Ziele und Funktionen für Noten übernommen bzw. favorisiert, die eigentlich so nicht gewollt sind (etwa die starke Betonung der Auslesefunktion; vgl. hierzu Ingenkamp 1975, Meister 1976).

Zusammenfassend kann also gefragt werden, inwieweit das Benotungsverhalten von Lehrern von deren individuellen pädagogischen Zielsetzungen abhängt. Zur empirischen Prüfung dieser Frage greifen wir auf ein einfaches instrumentalitätstheoretisches Modell aus der Handlungstheorie zurück (vgl. etwa Mitchell, Biglan 1971; Jungermann 1976), das Handlungspräferenzen durch (a) Handlungsziele und ihre Bewertung durch das Handlungssubjekt (Valenzen) und (b) Handlungs-Ergebnis-Erwartungen (Instrumentalitäten), also die mit einer Handlung verbundenen Erwartungen, daß durch diese Handlung die Handlungsziele erreicht oder behindert werden, erklärt. In einer ersten empirischen Prüfung dieses Modells für die Erklärung und Vorhersage von pädagogischen Handlungspräferenzen in erzieherischen Problemsituationen (keine Benotungssituationen) konnte seine Tauglichkeit belegt werden (siehe Krampen, Brandtstädter 1978). Solche entscheidungstheoretischen Modelle wurden bislang nur für das Entscheidungsverhalten von *Schülern*, die (im Fragebogenmedium) die Wahl zwischen verschiedenen Notenkombinationen hatten, und zur Abschätzung der subjektiven Nützlichkeit von bestimmten Noten bei den Schülern verwendet (vgl. Becker, Siegel 1958; Schmidt, Wiesenberg, Scholz 1971; Hommers 1977). Hier soll nun das Benotungsverhalten (bzw. die Präferenzen für die Vergabe bestimmter Noten) von *Lehrern* entscheidungstheoretisch analysiert werden. Neben einer Lehrerstichprobe wurden zusätzlich Lehramtskandidaten befragt, da sich in der Literatur ein Hinweis darauf findet, daß Studenten im allgemeinen etwas bessere Noten vergeben als berufserfahrene Lehrer (siehe Bonniol, Caverni, Noizet 1972a).

2. Fragestellungen

Die Kernfrage der vorliegenden Arbeit (Fragestellung 1) lautet:

1. Können die Präferenzen zur Vergabe von bestimmten Zeugnisnoten in Problemsituationen durch individuell vertretene Erziehungsziele und Instrumentalitätserwartungen der Lehrpersonen vorhergesagt werden?

Daneben ergeben sich eine Reihe von anderen Fragestellungen, die, bedingt durch die zu verwendende Methode, ebenfalls behandelt werden:

2. Welches Gewicht geben Lehrer und Lehramtskandidaten verschiedenen pädagogischen Zielsetzungen eines Zielkataloges? Gibt es in den Zielbewertungen Unterschiede zwischen Lehrern und Lehramtskandidaten?
3. Wie nehmen Lehrer und Lehramtskandidaten vorgegebene (problematische) Benotungssituationen wahr? Unterscheiden sich die beiden Stichproben in ihren Wahrnehmungen?

4. Welche Notengebung präferieren Lehrer und Lehramtskandidaten in den vorgegebenen Situationen? Wie hoch ist die Urteilskonkordanz innerhalb und zwischen den beiden Stichproben?
5. Gibt es Unterschiede in der Tauglichkeit des Instrumentalitäts-theoretischen Modells zur Vorhersage der Benotung zwischen Lehrern und Lehramtskandidaten?

3. Methode

Folgende Arbeitsschritte kennzeichnen den methodischen Ansatz der vorliegenden Untersuchungen und der Fragebogenkonstruktion:

1. Formulierung von pädagogischen Problemsituationen, die Benotungsverhalten verlangen, und von zugeordneten Handlungsalternativen:

Der Fragebogen umfaßte zwei Problemsituationen, in denen die Informanden Benotungen vornehmen sollten. Die Situationen wurden so konstruiert, daß die einfache Berechnung des arithmetischen Mittelwerts zu keiner eindeutigen Note führt. Zudem wurde in den Situationen zwischen schriftlichen und mündlichen Leistungen eines Schülers unterschieden, was evtl. zu zusätzlichen Bewertungsproblemen führen kann. Die Situationen lauteten:

Situation A: Ein Schüler hat in einem Schulfach im letzten halben Jahr die folgenden schriftlichen Leistungen gezeigt: 1 x sehr gut, 4 x gut, 1 x befriedigend. In mündlichen Prüfungen erreichte er die folgenden Noten: 1 x ausreichend, 4 x mangelhaft, 1 x ungenügend. Seine Beteiligung am Unterricht läßt zudem zu wünschen übrig.

Situation B: Ein sprachgestörter Schüler hat . . . (sonst wie Situation A).

Die Informanden sollten an Hand von 7stufigen Schätzskalen die beiden Problemsituationen beurteilen. Sie sollten angeben, ob die jeweilige Situation (1) klar durch Vorschriften geregelt versus offiziell nicht geregelt ist, (2) für sie einfach versus kompliziert ist, (3) ärgerlich versus nicht ärgerlich ist, (4) subjektiv gut bekannt versus nicht bekannt ist und (5) sehr häufig versus äußerst selten im Schulalltag vorkommt.

Als Handlungsalternativen war die Vergabe der Notenstufen „sehr gut“, „gut“, „befriedigend“, „ausreichend“ und „mangelhaft“ für die Note in dem Schulfach auf dem Halbjahreszeugnis (dem ja eine besondere pädagogische Bedeutung zukommen sollte) vorgegeben.

2. Ermittlung der faktischen Präferenz des Individuums i für die Handlungsalternative h (P_{ih}) in den Situationen:

Die Informanden wurden gebeten, sich in die Situationen hineinzuversetzen, die vorgegebenen Notenstufen mit ihren eigenen Benotungen in der jeweiligen Situation zu vergleichen und sie dementsprechend in eine Präferenzrangreihe zu bringen.

3. Formulierung eines Kataloges von pädagogischen Zielsetzungen und Ermittlung der persönlichen Bedeutung bzw. Valenz, die die Person i der Zielsetzung z zuschreibt (V_{iz}):

Auf der Grundlage früherer Arbeiten zu Erziehungs- und Rollenzielen von Lehrern (Krampen 1979) wurden die folgenden zehn Zielsetzungen ausgewählt und den Informanden zur Bewertung vorgelegt: (1) als Lehrer sollte man den Ehrgeiz des Schülers herausfordern; (2) als Lehrer sollte man die Kreativität des Schülers fördern; (3) als Lehrer sollte man das Selbstbewußtsein des Schülers fördern; (4) als Lehrer sollte man

den Schüler zu Disziplin und Ordnungsliebe anhalten; (5) als Lehrer sollte man auf sprachliche Leistungen nicht so großen Wert legen; (6) als Lehrer sollte man beim Schüler beliebt sein; (7) als Lehrer sollte man dem Schüler gegenüber Autorität wahren; (8) als Lehrer sollte man Wissen und Bildung vermitteln; (9) als Lehrer sollte man sich um absolute Objektivität bemühen; (10) als Lehrer sollte man ein vertrauensvolles Verhältnis zum Schüler haben. Diese Zielsetzungen sollten zunächst allgemein und dann speziell für den Fall von Schülern mit Sprachstörungen bewertet werden.

Die Skalierung der Zielsetzungen wurde analog zum Vorgehen in der (normativen) multiattributen Nutzentheorie (vgl. etwa Edwards, Guttentag 1975) durchgeführt. Die Informanden sollten die zehn Zielsetzungen zunächst nach dem Kriterium der persönlichen Wichtigkeit und Bedeutung in eine Rangreihe bringen. Dem Ziel, das die Person auf Rangplatz 10 gesetzt hat, sollten dann 10 Wertpunkte zugeteilt werden. Durch einfache Verhältnisschätzungen sollten dann auf die folgenden Zielsetzungen jeweils so viele Wertpunkte verteilt werden, wie sie für die Person subjektiv wichtiger und bedeutsamer gegenüber der jeweils unwichtigeren Zielsetzung sind. Um eine Vergleichbarkeit der individuellen Zielbewertungen herzustellen, wurde für jede Person die Summe aller verteilten Wertpunkte gebildet und die auf die einzelnen Zielsetzungen entfallenden Wertpunkte durch diese Summe dividiert (zu diesem Vorgehen siehe auch Edwards, Guttentag 1975).

4. Ermittlung der Instrumentalität, die die Person i der Handlungsalternative h (Notenstufe) für die Erreichung der Zielsetzung z zuschreibt (I_{ihz}):

Die Informanden sollten für jede Situation die Instrumentalität der Reaktionen zur Erreichung der vorgegebenen zehn Zielsetzungen auf einer Skala abschätzen. Die Skala reichte von +10 („die Note trägt in starkem Maße zur Zielerreichung bei“) über 0 („ist ohne Bedeutung für die Zielerreichung“) bis -10 („die Note verhindert in starkem Maße die Zielerreichung“). Dieser Arbeitsschritt war für die Informanden der aufwendigste; hier waren pro Situation $10 \times 5 = 50$ Einzelschätzungen abzugeben.

5. Berechnung eines theoretischen Präferenzwertes zur Vorhersage der Präferenz der Person i für die Handlungsalternative h aus den in (3) und (4) ermittelten Valenz- und Instrumentalitätswerten nach der Formel

$$\hat{P}_{ih} = \sum_z V_{iz} I_{ihz}$$

die dem instrumentalitätstheoretischen Ansatz (SEU-Modell; vgl. etwa Jungermann 1976) entspricht.

6. Vergleich der faktischen mit den theoretischen Präferenzwerten (Fragestellung 1).

Ein Fragebogen, der die Arbeitsschritte (1) bis (4) umfaßte, wurde insgesamt 51 Grund- und Hauptschullehrern ($\bar{x} = 29;5$ Jahre, $s = 8;4$ Jahre, $N = 30$ Lehrerinnen) und 87 Lehramtskandidaten ($\bar{x} = 20;4$ Jahre, $s = 1;8$ Jahre, $N = 70$ Studentinnen) für das Lehramt an Grund- bzw. Hauptschulen aus dem süddeutschen Raum vorgelegt. Die Berufserfahrung der befragten Lehrer betrug im Durchschnitt $5;10$ Jahre ($s = 8;3$ Jahre). Die Lehramtskandidaten hatten Unterrichtserfahrungen aus studienbegleitenden Praktika und Blockpraktika während der Semesterferien.

Tabelle 1. Mittelwerte, Standardabweichungen und Rangplätze der Zielbewertungen für den allgemeinen und für den Fall von sprachgestörten Schülern in beiden Stichproben

Zielsetzung ^a	Lehrer (N = 51)						Lehramtskandidaten (N = 87)					
	allgemein			mit Sprachstörung			allgemein			mit Sprachstörung		
	\bar{x}	s	Rang	\bar{x}	s	Rang	\bar{x}	s	Rang	\bar{x}	s	Rang
1. Ehrgeiz	6,59	9,89	6	7,96	10,22	4	6,12	6,42	6	8,93	9,07	4
2. Kreativität	16,49	13,07	2	13,57	12,26	3	17,24	12,34	2	12,49	7,79	3
3. Selbstbew.	13,96	14,23	3	19,41	16,18	2	18,78	17,47	1	24,59	18,77	1
4. Disziplin	4,77	4,42	8	3,82	4,41	9	3,03	3,73	9	2,75	4,04	10
5. Sprachl. L.	2,77	3,77	10	6,69	7,17	7	2,64	3,14	10	8,44	9,14	5
6. Beliebtheit	5,26	4,98	7	5,92	5,28	8	5,47	5,72	7	5,81	6,26	8
7. Autorität	4,31	4,02	9	3,45	4,01	10	3,59	5,16	8	2,81	5,14	9
8. Wissen	11,98	8,42	4	7,59	6,93	5	11,39	11,88	4	6,78	8,17	7
9. Objektivität	9,77	9,53	5	7,12	9,88	6	10,69	8,46	5	7,40	8,25	6
10. vert. Verh.	21,04	15,35	1	20,12	15,45	1	17,22	14,94	3	16,00	13,82	2

^a genaue Bezeichnung der Zielsetzungen im Text

4. Ergebnisse

4.1. Erziehungs- und Rollenziele

In Tabelle 1 sind die Mittelwerte, Standardabweichungen und Rangplätze der Zielbewertungen in der Stichprobe der Lehrer und der Lehramtskandidaten für den allgemeinen Fall und für den Fall der sprachgestörten Schüler zusammengestellt (Fragestellung 2). Allgemein ist festzustellen, daß Schülerorientierte Entwicklungsziele (das sind die Ziele 1 bis 4 und das Ziel 10) im Durchschnitt höhere Rangplätze einnehmen als die Rollenziele für das eigene Lehrerverhalten (das sind die Ziele 5 bis 9). Dies gilt vor allem für die Zielsetzungen „vertrauensvolles Verhältnis zum Schüler“, „Selbstbewußtsein“ und „Kreativität“. Dies bestätigt die Befunde von Krampen (1979), wobei natürlich zu beachten ist, daß Zielkataloge immer nur relative Aussagen über die Wichtigkeit und Bedeutung von einzelnen pädagogischen Zielsetzungen zulassen.

Vergleicht man die Zielbewertungen von Lehrern und Lehramtskandidaten (mit dem Mann-Whitney U-Test; vgl. etwa Lienert 1973²), so ergeben sich bei zwei Zielsetzungen statistisch signifikante Differenzen. Das Schülerorientierte Erziehungsziel „Selbstbewußtsein des Schülers fördern“ wird von Studenten sowohl für den allgemeinen als auch für den speziellen Fall der Schüler mit Sprachstörungen höher bewertet ($p < .01$), das Erziehungsziel „vertrauensvolles Verhältnis zum Schüler“ wird von ihnen (ebenfalls in beiden Fällen) geringer gewichtet als von Lehrern ($p < .01$).

Mittelwertsvergleiche der Zielbewertungen für den allgemeinen Fall und für den Fall von sprachgestörten Schülern innerhalb einer Stichprobe (durchgeführt mit dem Wilcoxon-Test; vgl. Lienert 1973²) ergaben folgende signifikante Unterschiede: (1) Lehrer bewerten die Zielsetzungen „Selbstbewußtsein“, „auf Sprache geringeren Wert legen“ (jeweils $p < .01$) und „Ehrgeiz“ ($p < .05$) bei Schülern mit Sprachstörungen höher, die Erziehungsziele „Kreativität“, „Wissen und Bildung“ und „Objektivität“ (jeweils $p < .01$) geringer als bei Schülern ohne Sprachstörungen; (2) Bei Lehramtskandidaten ergibt sich das gleiche Bild, wobei nur festzuhalten ist, daß bei ihnen die Differenz bei dem Ziel „Ehrgeiz“ auf dem 1 %-Niveau signifikant ist. Eine Inspektion der Mittelwertsunterschiede zeigt allgemein, daß die Differenzen bei Lehramtskandidaten etwas stärker ausgeprägt sind als bei den berufserfahrenen Lehrern.

Für die Zielbewertungen wurden (getrennt für beide Stichproben jeweils für den allgemeinen und für den Fall der sprachgestörten Schüler) vier Hauptachsen-Faktorenanalysen (R^2 als Startkommunalität, Varimax-Rotation) berechnet. In allen Analysen hatten jeweils vier Faktoren Eigenwerte $e_j > 1.00$; diese Faktoren erklären zwischen 63,9 % und 70,3 % der Gesamtvarianzen. Auf eine ausführliche Darstellung und Interpretation dieser dimensionsanalytischen Ergebnisse soll hier verzichtet werden, da sie nur „am Rande“ zu unserer Fragestellung gehören. Es sei lediglich vermerkt, daß

sich in allen Analysen jeweils neben spezifischen, Schüler-orientierten Ziel-dimensionen zwei Faktoren identifizieren lassen, die inhaltlich ähnlich sind zu den Faktoren „Autoritäts- und Disziplinentorientierung“ und „rational-systematische Orientierung“ bei Krampen (1979).

3. 2. Wahrnehmung der vorgegebenen Problemsituationen

Die Mittelwerte und Standardabweichungen der Situationswahrnehmungen durch beide Stichproben sind in Tabelle 2 aufgeführt (Fragestellung 3). Mittelwertsvergleiche, durchgeführt mit Mann-Whitney U-Tests, zeigen, daß Lehr-amtskandidaten Situation A und Situation B als komplizierter und als ärgerlicher erleben ($p < .01$) als die Lehrer, gleichzeitig geben sie auch an, daß

Tabelle 2. Wahrnehmung der Benotungssituationen durch Lehrer und Lehramtskandi-daten.

Attribut	Lehrer (N = 51)				Kandidaten (N = 87)			
	Situation A		Situation B		Situation A		Situation B	
	\bar{x}	s	\bar{x}	s	\bar{x}	s	\bar{x}	s
1. (nicht) durch Vor-schriften geregelt	4,10	1,94	4,96	1,73	3,80	1,74	4,75	1,68
2. einfach versus kompliziert	4,62	1,50	5,02	1,62	5,45	1,41	5,54	1,64
3. ärgerlich versus nicht ärgerlich	4,18	1,92	4,71	2,03	3,63	1,97	4,04	2,06
4. bekannt versus unbekannt	4,31	2,07	4,63	1,92	2,89	1,88	4,00	2,02
5. häufig versus selten	4,88	1,27	4,51	1,56	3,26	1,41	3,37	1,47

ihnen die Situationen bekannter sind (!), und daß sie glauben, daß sie im Schulalltag häufiger auftreten ($p < .01$). Situationsvergleiche innerhalb der Stichproben (Wilcoxon-Test) ergaben, daß sowohl Lehrer als auch Lehramts-kandidaten die Situation B (Schüler mit Sprachstörungen) als weniger durch Vorschriften geregelt, als komplizierter und als weniger ärgerlich erleben als die Situation A (alle $p < .01$). Bei Lehramtskandidaten ergibt sich zusätzlich, daß sie Situation A im Durchschnitt als „bekannter“ einschätzen als Situation B ($p < .01$).

3. 3. Notengebung in den Situationen (Handlungspräferenzen)

Die Mediane der faktischen Präferenzen für die Notenstufung in den Situa-tionen A und B sind getrennt für beide Stichproben in Tabelle 3 aufgeführt (Fragestellung 4). Friedman-Varianzanalysen für Ordinaldaten bestätigen, daß

Tabelle 3. Friedman-Varianzanalysen der direkt erhobenen Benotungs-Präferenzen in den beiden Stichproben für Situation A und B

Stichprobe	Situation	Mittlerer Rangplatz (Median) der Note					Chi ² Ränge	W
		1	2	3	4	5		
Lehrer	A	4,20	2,32	1,22	2,64	4,69	175,50 +	.70 +
Lehrer	B	3,79	1,32	1,73	3,23	4,84	165,00 +	.66 +
Studenten	A	4,17	2,08	1,12	2,93	4,71	327,66 +	.76 +
Studenten	B	3,59	1,24	1,82	3,44	4,87	330,67 +	.77 +

+ p < .001

die einzelnen Notenstufen in signifikanter Weise unterschiedlich bevorzugt werden (siehe Tabelle 3). Für den Fall des Schülers mit Sprachstörungen (Situation B) setzten beide Stichproben im Mittel die Note „gut“ auf den ersten Rangplatz, für den allgemeinen Fall (Situation A) die Note „befriedigend“. Die Notenstufen „ausreichend“, „sehr gut“ und „mangelhaft“ erhalten die Rangplätze 3, 4 und 5. Nur als Tendenz ist erkennbar, daß Studenten etwas besser bewerten als die Lehrerstichprobe (vgl. hierzu Bonniol, Carverni, Noizet 1972b); berechnete Mittelwertvergleiche (Mann-Whitney U-Tests) ergaben keine Signifikanzen. Neben den Unterschieden in der Bevorzugung der einzelnen Notenstufen sind die ebenfalls in Tabelle 3 mitgeteilten Konkordanzkoeffizienten W von Interesse. Ihr durchgängig hoher Wert belegt die hohe interindividuelle Urteilsübereinstimmung der befragten Personen in den Rangreihen der Notenpräferenzen.

Auch für die modellspezifisch berechneten (theoretischen) Präferenzwerte wurden Rangdatenvarianzanalysen berechnet. Sie bestätigen im wesentlichen die Befunde zu den direkt erhobenen Präferenzwerten. Es zeigen sich jedoch einige Vertauschungen in den Rangreihen der Mediane, die darauf hindeuten, daß die Übereinstimmung von faktischen und theoretischen Präferenzrangreihen in der vorliegenden Untersuchung nicht so gut sein werden wie bei Krampen und Brandtstädter (1978). Die Konkordanzkoeffizienten W für die modellspezifisch ermittelten Präferenzrangreihen liegen ebenso wie bei Krampen und Brandtstädter (1978) in der Lehrerstichprobe etwas höher als die für die faktischen Rangreihen (Lehrerstichprobe: Situation A W = .73, Situation B W = .84); für die Stichprobe der Lehramtskandidaten zeigt sich dieses Bild nicht (Situation A: W = .70, Situation B: W = .77).

3. 4. Vergleich der direkt erfragten mit den modellspezifisch berechneten Reaktionspräferenzen

Zum Vergleich der faktischen mit den theoretischen Präferenzwerten muß das Ausmaß der Beziehung zwischen den Rangordnungen der Reaktionsalter-

nativen für jede Situation bestimmt werden (Fragestellung 1). Dazu wird die Kendall-Summe S verwendet, die durch das Auszählen der Proversionen und Inversionen in den Rangreihen gebildet wird. Dieses Differenzmaß für zwei Rangordnungen beruht also auf der Fehlerordnung der Ränge innerhalb der Vergleichsreihe. Für jedes Individuum wird für jede Situation der Inversionswert T , der in einer eindeutigen Beziehung zu S steht, berechnet. Dabei entspricht ein T von 0 der völligen Übereinstimmung beider Präferenzrangreihen ($\tau = 1.00$), ein T von 5 belegt, daß keine Beziehungen zwischen beiden Rangreihen besteht ($\tau = .00$) und ein T von 10 bedeutet, daß beide Rangreihen sich völlig unterscheiden, also gegensinnig verlaufen ($\tau = -1.00$).

In Tabelle 4 sind die Häufigkeitsverteilungen der T -Werte in den Situationen A und B zusammenfassend für die Stichprobe der Lehrer und der der Lehramtskandidaten dargestellt. Diese Zusammenfassung beider Stichproben ist hier möglich, da ein Kolmogoroff-Smirnov-Omnibus-Test (vgl. Lienert 1973²) keine Verteilungsunterschiede zwischen Lehrern und Studenten aufdeckte (Fragestellung 5). Die allgemeine explanative und prognostische Tauglichkeit des Modells, nämlich ob die erhaltenen Häufigkeitsverteilungen der T -Werte signifikant von der Häufigkeitsverteilung abweicht, die unter der Nullhypothese (kein Zusammenhang zwischen faktischen und theoretischen Präferenzen) erwartet wird, wurden mit Kolmogoroff-Smirnov-Anpassungstests (KSA-Test; vgl. Lienert 1973²) überprüft. Bei der Bestimmung der unter der Nullhypothese zu erwartenden Häufigkeitsverteilung wurde die für die Kendall-Summe S bei Kendall (1948) tabellierte Prüfverteilung verwendet (siehe Tabelle 4: f_e).

Zur statistischen Prüfung wurde der KSA-Test dem χ^2 -Test vorgezogen, da in ihn als ein nichtparametrischer Test geringere meßtheoretische Voraussetzungen eingehen, und da er zudem als der statistisch strengere Test bezeichnet wird (Siegel 1956).

Die höchste Abweichung der empirischen T -Werteverteilung von der erwarteten Verteilung liegt mit $K^0 = 0,288$ für Situation A über der Prüfgröße ($K^0 = 0,1387$ für $df = 138$ und $\alpha = .01$); die Abweichung ist also statistisch überzufällig hoch. Der mittlere Rangkorrelationskoeffizient zwischen der faktischen und theoretischen Rangreihe beträgt allerdings nur $\tau = .19$ (signifikant bei $p < .001$; Signifikanzbestimmung über die z -Verteilung nach Siegel 1956), was den signifikanten Zusammenhang der beiden Präferenzrangreihen in Situation A bestätigt, zugleich aber die geringe Höhe des Zusammenhanges vor Augen führt.

Auch in Situation B ist die Abweichung der empirischen T -Werteverteilung von der Erwarteten mit $K^0 = 0,440$ (siehe Tabelle 4) auf dem 1 %-Niveau signifikant. Der mittlere Rangkorrelationskoeffizient beträgt $\tau = .41$ ($p < .001$), was auf die höhere Präzision des instrumentalitätstheoretischen Modells in Situation B (sprachgestörter Schüler) hindeutet. Für die Berechnung der theoretischen Präferenzwerte wurden hier die für Schüler mit Sprachstörungen spezifischen Zielbewertungen, also nicht nur andere Instru-

Tabelle 4. Kolmogoroff-Smirnov-Anpassungstests der empirischen T-Wertverteilungen an die unter der Nullhypothese erwartete Verteilung (N = 138) ^a

T-Wert	f_e	$P(f_e)_{cum}$	Situation A		Situation B	
			f_b	$P(f_e)_{cum} - P(f_b)_{cum}$	f_b	$P(f_e)_{cum} - P(f_b)_{cum}$
0	1,15	0,008	5	0,028	13	0,086
1	4,65	0,042	2	0,009	9	0,117
2	10,35	0,117	9	0,001	33	0,282
3	17,35	0,242	20	0,019	31	0,381
4	22,90	0,408	60	0,288	31	0,440
5	25,40	0,592	24	0,277	13	0,350
6	22,90	0,758	12	0,199	6	0,228
7	17,25	0,883	2	0,088	2	0,117
8	10,35	0,958	4	0,042	0	0,042
9	4,65	0,992	0	0,008	0	0,008
10	1,15	1,000	0	0,000	0	0,000
N	138,00		138		138	

^a f_e = erwartete Häufigkeiten; f_b = beobachtete Häufigkeiten

mentalitätsschätzungen, sondern auch andere Valenzen als in Situation A, verwendet. Bei Krampen und Brandtstädter (1978) wurden dagegen die gleichen Zielvalenzen mit situationsspezifischen Instrumentalitäten verknüpft. Dieses Vorgehen führte hier für die Situation B (also allgemeine Valenzen und situationsspezifische Instrumentalitäten) zu einer T-Werte-Verteilung, die zwischen den in Tabelle 4 für die Situation A und B angegebenen liegt.

5. Diskussion

Reaktionspräferenzen für die Vergabe von Zeugnisnoten in pädagogischen Problemsituationen können also partiell durch ein instrumentalitätstheoretisches Modell, das auf individuell vertretene Erziehungsziele und mit der Notenvergabe verbundene Instrumentalitätserwartungen rekurriert, erklärt und vorhergesagt werden. Das verwendete entscheidungstheoretische Modell vermag einen Beitrag zur Analyse des Benotungsverhaltens (nicht seiner Bewertung!) zu leisten. In Folgestudien sollten auch Modellerweiterungen beachtet werden, wie sie allgemein etwa von Fishbein und Ajzen (1975; Rollenerwartungen, Rollenzwänge für das Benotungsverhalten) vorgeschlagen wurden. Es bleibt zunächst festzuhalten, daß die Präzision des Modells in der vorliegenden Arbeit, wo es zur Analyse von Benotungsverhalten eingesetzt wurde, geringer ist als in der Arbeit von Krampen und Brandtstädter (1978), wo es zur Erklärung und Vorhersage von pädagogischen Handlungspräferenzen in allgemeinen, nicht mit Fragen der Zensierung verbundenen erzieherischen Problemsituationen verwendet wurde. Als Ursache dafür kann festgestellt werden, daß für die Handlungsalternativen „sehr gut oder gut als Zeugnisnote geben“ die nach der modellspezifischen Formel bezeichneten Produktsummen von Valenzen und Instrumentalitäten relativ hohe Werte einnehmen, also nach dem Modell eher präferiert werden als in den faktischen, direkt erhobenen Präferenzen. Eine Feinanalyse des Datenmaterials zeigt, daß dies vor allem durch die sehr hoch bewerteten Zielsetzungen „vertrauensvolles Verhältnis zum Schüler“, „Selbstbewußtsein des Schülers fördern“ und „Ehrgeiz des Schülers herausfordern“ zustandekommt, für die diese Handlungsalternativen auch relativ hohe positive Instrumentalitätsschätzungen einnehmen; das Ergebnis sind hohe Produktsummen für diese Handlungen, die die Abweichungen der theoretischen von der faktischen Präferenzrangreihe, und somit die höheren T-Werte, bedingen. Die Zielsetzungen, bei denen negative Instrumentalitätsschätzungen für die Vergabe guter Noten (wie etwa bei den Zielsetzungen „Objektivität“, „Wissen und Bildung vermitteln“) vorliegen, verfügen dagegen über geringere Valenzen. Ihre handlungsleitende Bedeutung ist nach dem verwendeten Modell geringer, da sie die Produktsummen in schwächerem Ausmaß beeinflussen als die höher bewerteten Ziele, die gleichzeitige hohe positive Instrumentalitäten bei der Vergabe guter Noten haben.

Diese Ergebnisse legen nahe, daß es die befragten Lehrer und Lehramtskandidaten nur in relativ geringem Maße für möglich halten, durch die Vergabe von bestimmten Zeugnisnoten von ihnen selbst hoch bewertete Erziehungsziele zu erreichen, und daß die Erreichung von Erziehungszielen wie „Selbstbewußtsein des Schülers fördern“, „vertrauensvolles Verhältnis zum Schüler haben“, „Ehrgeiz des Schülers herausfordern“ etc., wenn überhaupt, dann nur durch die Vergabe *guter* oder *sehr guter* Noten möglich wäre. Die reale pädagogische Effektivität von Noten steht also bei den befragten berufserfahrenen und in Ausbildung stehenden Lehrern in Zweifel, was ähnliche Befunde auf der Schülerseite (vgl. etwa Pfeiffer 1977, Jopt 1978) bestätigt.

Sehr gute Noten können aus schulrechtlichen Gründen und wegen der fachlichen, auf Lehrinhalte gerichteten Aufgaben des Lehrers in den vorgegebenen Problemsituationen nicht vergeben werden. Dies wird bei den faktischen Reaktionspräferenzen deutlich. Der dennoch bestehende Handlungsspielraum des Lehrers bei der Notengebung zeigt sich in der vorliegenden Untersuchung darin, daß Schüler mit Sprachstörungen bei gleichen Einzelleistungen eher eine bessere Zeugnisnote gegeben würde als Schülern ohne Sprachstörungen. Hier deutet sich eine „Chancenausgleichsfunktion“ (Zielinski 1974), die Zensuren prinzipiell haben können, an. In Situation B, wo die Problematik der Zensurengebung durch die Sprachstörung des zu benotenden Schülers verschärft war, erwies sich ja auch das instrumentalitätstheoretische Model als präziser. Dies bestätigt die These, daß entscheidungstheoretische Ansätze gerade in Situationen von Bedeutung sind, die für das Entscheidungssubjekt in hohem Maße problematisch sind. Zugleich ist es ein Hinweis darauf, daß in solchen Situationen wohl eher Reflektionen über die mit der Notengebung verfolgten pädagogischen Zielsetzungen und über die möglichen Nebeneffekte ablaufen.

Durch die vorgelegten Befunde wird nun nicht nur die „pädagogische“ Kritik an Noten (vgl. etwa Ingenkamp 1975, Meister 1976) verstärkt, es ergeben sich auch Schlußfolgerungen für die Aus- und Weiterbildung von Lehrern: Trainingsseminare zur Notengebung und Schülerbeurteilung sollten sich nicht nur auf methodologische Aspekte wie Objektivität, Vergleichbarkeit, Reliabilität etc. von Zensuren konzentrieren (vgl. etwa Bonniol et al. 1972b, Feldmesser 1972), was letztlich auf eine Schulung der Teilnehmer in „Beobachtungsmethoden“ hinausläuft, sondern auch auf die pädagogischen Zielsetzungen und die mit der Vergabe bestimmter Noten in speziellen Situationen verbundenen Handlungsziele eingehen. In Zusammenhang mit Überlegungen zu impliziten Persönlichkeitstheorien, Kausalattributionen, Bezugsnormorientierungen etc. können solche entscheidungstheoretischen Analysen etwa in Form von Fallstudien durchgeführt werden.

Abschließend soll noch kurz auf die Ergebnisse zu den Fragestellungen (2) bis (4) eingegangen werden. Die Befunde zu den Erziehungs- und Rollenzielen bestätigen im wesentlichen vorliegende Ergebnisse (Krampen 1979). Die verwendete, von der multiattributen Nutzentheorie übernommene Skalierung

der Zielsetzungen erwies sich als für die Informanden zu aufwendig und erbrachte gegenüber der einfachen Wertpunkteverteilung (siehe Krampen, Brandtstädter 1978) bei der deskriptiven Verwendung der Instrumentalitäts-theorie keine Vorteile (anders ist dies bei der normativen Verwendung; vgl. hierzu Edwards, Guttentag 1975; Kasubek, Aschenbrenner 1978). Die Zielbewertungen von Lehrern und Lehramtskandidaten unterscheiden sich nur gering, dagegen liegen bemerkenswerte Differenzen in den Zielgewichtungen für den allgemeinen und den Fall von sprachgestörten Schülern vor, was die Globalität von Erziehungszielvorstellungen (Lukesch 1975) in Frage stellt. Der Befund, daß Studenten dazu tendieren, bessere Noten zu geben als berufserfahrene Lehrer (Bonniol et al. 1972a), konnte hier nicht bestätigt werden. Unterschiede zwischen Lehramtskandidaten und Lehrern in der Wahrnehmung von problematischen Benotungssituationen wurden ermittelt; sie sind auf der Basis der jeweiligen Erfahrungshintergründe gut zu interpretieren. Generationsbedingte Differenzen können hier zwar nicht ausgeschlossen werden (dazu wären Längsschnittuntersuchungen notwendig), sind jedoch unwahrscheinlich.

Anmerkung

- 1 Zielinski (1974) berichtet nur anekdotisch über eine Umfrage zu den Funktionen von schulischen Zensuren bei Lehrern. Eine umfassende Funktionsanalyse, die neben der Lehrersicht etwa auch die Schülersichtweise beachtet, steht bislang nach unserem Wissen noch aus. Empirische Untersuchungen dazu sind von uns in Arbeit.

Literatur

- Becker, S.W., Siegel, S.: Utility of grades — Level of aspiration in a decision theory context. *J. Exper. Psychol.* 55, 81–85 (1958)
- Bonniol, J.-J., Caverni, J.-P., Noizet, G.: Scholastic status of pupils as a determinant of marks assigned on a specific task. *Cahiers de Psychologie* 15, 83–92 (1972a)
- Bonniol, J.-J., Caverni, J.-P., Noizet, G.: Experimental training in evaluating scholastic tasks. *Cahiers de Psychologie* 15, 93–104 (1972b)
- Edwards, W., Guttentag, M.: Experiments and evaluation — A reexamination. In: Bennett, C.A., Lumsdale, A.A. (Ed.), *Evaluation and experiment*. New York 1975, p. 409–463
- Eells, W.C.: Reliability of repeated grading of essay type examinations. *J. Educat. Psychol.* 27, 48–52 (1930) (Deutsche Übersetzung in: Ingenkamp, K. (Ed.), *Die Fragwürdigkeit der Zensurengebung*. Weinheim: Beltz 1971, p. 117–122)
- Feldmesser, R.A.: The positive functions of grades. *Educat. Rec.* 53, 66–72 (1972)
- Ferdinand, W.: Das Vorurteil des Lehrers über die Leistungsfähigkeit bestimmter Schüler im Spiegel der Aufsatzzensur. *Schule u. Psychol.* 18, 92–95 (1971)
- Finalyson, D.S.: The reliability of the marking of essays. *British J. Educat. Psychol.* 27, 126–134 (1951)
- Fishbein, M., Ajzen, I.: *Belief, attitude, intention and behavior*. Reading, Mass.: Addison-Wesley 1975

- Hadley, S.T.: A school mark — Fact or fancy. *Educational Administration and Supervision* 40, 305–312 (1954) (Deutsche Übersetzung in: Ingenkamp, K. (Ed.), *Die Fragwürdigkeit der Zensurengebung*. Weinheim: Beltz 1971, p. 134–141)
- Hanke, B., Lohmöller, J.-B., Mandl, H.: Schülerbeurteilung. Schicht und Schullaufbahn. *Z. Entw. Psychol. u. Pädagog. Psychol.* 7, 113–133 (1970)
- Heller, K.: Zur Problematik der Leistungsbeurteilung in der Schule. *Psychol. in Erziehung und Unterricht* 21, 105–124 (1974)
- Herrig, D., Krampen, G.: Wie erklären Schüler und deren Lehrer das Zustandekommen von Zeugnisnoten? *Psychologie in Unterricht und Erziehung* 26, 291–294 (1979)
- Hofer, M.: Die Schülerpersönlichkeit im Urteil des Lehrers. Weinheim: Beltz 1974 (3. Auflage)
- Höger, D.: Allgemeine Voraussetzungen für Leistungsmessung und Leistungsbeurteilung. In: Stephan, E., Schmidt, W. (Ed.), *Messen und Beurteilen von Schülerleistungen*. München: Urban & Schwarzenberg 1978, p. 56–76
- Hommers, W.: Zur Skalierung des subjektiven Werts von Zensuren bei Schulkindern. *Z. f. Psychol.* 785, 423–433 (1977)
- Ingenkamp, K.H. (Ed.): *Die Fragwürdigkeit der Zensurengebung*. Weinheim: Beltz 1971
- Ingenkamp, K.H.: *Pädagogische Diagnostik*. Weinheim: Beltz 1975
- Jopt, U.-J.: Wenn Schüler Zensuren geben könnten — Attributive Begleitprozesse diskrepanter Leistungsevaluationen. Vortrag, gehalten auf dem 31. Kongreß der Deutschen Gesellschaft für Psychologie in Mannheim 1978
- Jungermann, H.: *Rationale Entscheidungen*. Bern: Huber 1976
- Kasubek, W., Aschenbrenner, M.: Optimierung subjektiver Urteile. Anwendung der multiattributen Nutzentheorie bei medizinischen Therapieentscheidungen. *Z. exper. u. angew. Psychol.* 15, 594–616 (1978)
- Kendall, E.W.: *Rank correlation methods*. London: Griffin 1948
- Kleber, E.W.: Die Beurteilerfunktion des Lehrers. In: Kleber, E.W., Meister, H., Schwarzer, C., Schwarzer, R. (Ed.), *Beurteilung und Beurteilungsprobleme*. Weinheim: Beltz 1976, p. 17–38
- Kleiter, E.: Über Theorie und Modell kategorialer Fehler des Lehrerurteils. *Psychol. Beitr.* 15, 185–229 (1973)
- Knapp, A.: Der Einfluß der Fachleistungsdifferenzierung auf Persönlichkeitsveränderungen von Schülern im sozialen und emotionalen Bereich. *Psychol. in Erziehung u. Unterricht* 25, 306–314 (1978)
- Knoche, W.: *Jungen, Mädchen, Lehrer und Schulen im Zensurenvergleich*. Weinheim: Beltz 1969
- Krampen, G.: Erziehungsleitende Vorstellungen von Lehrern. *Z. exper. u. angew. Psychol.* 25, 94–112 (1979)
- Krampen, G., Brandstädter, J.: Instrumentalitätstheoretische Vorhersage pädagogischer Handlungspräferenzen. *Z. Entw. Psychol. u. Pädagog. Psychol.* 10, 8–17 (1978)
- Lienert, G.A.: *Verteilungsfreie Methoden in der Biostatistik*. Band 1. Meisenheim: Hain 1973
- Lukesch, H.: *Erziehungsstile*. Stuttgart: Kohlhammer 1975
- Mandl, H., Lohmöller, J.-B., Hanke, B.: Schichtspezifische Faktorenstrukturen in Schülerbeurteilungen? Ein Beitrag zur impliziten Persönlichkeitstheorie des Lehrers. *Unterrichtswissensch.* 2, 19–29 (1975)
- Masendorf, F., Tscherner, K., Tücke, M.: Clusteranalytisch ermittelte Beurteilungstendenzen bei der Einschätzung der Schülerpersönlichkeit durch den Lehrer. *Z. Entw. Psychol. u. Pädagog. Psychol.* 6, 19–27 (1947)
- Meister, H.: Der Beitrag einzeldiagnostischer Ansätze zur Bewältigung von Beurteilungsproblemen. In: Kleber, E.W., Meister, H., Schwarzer, C., Schwarzer, R. (Ed.), *Beurteilung und Beurteilungsprobleme*. Weinheim: Beltz 1976, p. 253–294
- Meyer, W.-U., Butzkamm, A.: Ursachenerklärung von Rechennoten. I. Lehrerattributierungen. *Z. Entw. Psychol. u. Pädagog. Psychol.* 7, 53–66 (1975)

- Mitchell, T.R., Biglan, A.: Instrumentality theories – Current uses in psychology. *Psychol. Bulletin* 76, 432–454 (1971)
- Pfeiffer, H.: Zeugnissnoten und ihre Bewertung durch Schüler des 4. Schuljahres. *Psychol. in Erziehung u. Unterricht* 24, 267–275 (1977)
- Rheinberg, F.: Bezugsnorm-Orientierung. Versuch einer Integration motivierungsbedeutender Unterschiede. In: Tack, W.H. (Ed.), Bericht über den 30. Kongreß der Deutschen Gesellschaft für Psychologie. Göttingen: Hogrefe 1976, p. 318–319
- Rheinberg, F., Schmalt, H.-D., Wasser, I.: Ein Lehrerunterschied, der etwas ausmacht. *Z. Entw. Psychol. u. Pädagog. Psychol.* 10, 3–7 (1978)
- Schäfer, E.: Eine Untersuchung zur Konstanz und Differenzierung des Stereotyps „schlechter Schüler“. *Z. exper. u. angew. Psychol.* 12, 94–112 (1975)
- Schmidt, H.-D., Wiesenberg, M., Scholz, O.B.: Über Bedingungen der Veränderung von Nutzenskalen der schulischen Leistungsbewertung. *Z. Psychol.* 179, 281–299 (1971)
- Siegel, S.: *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill 1956
- Steinkamp, G.: Lehrer voller Vorurteile? Soziologische Analyse der Schülerbeurteilung in der Volksschule. *Die Deutsche Schule* 60, 802–816 (1968)
- Stephan, E.: Leistungsmessung und Leistungsbeurteilung im Überblick. In: Stephan, E., Schmidt, W. (Ed.), *Messen und Beurteilen von Schülerleistungen*. München: Urban & Schwarzenberg 1978, p. 13–36
- Zielinski, W.: Die Beurteilung von Schülerleistungen. In: Weinert, F.E., Graumann, C.F., Heckhausen, H., Hofer, M. (Ed.), *Funk-Kolleg Pädagogische Psychologie*. Band 2. Frankfurt/M.: Fischer 1974, p. 879–900

Dipl.-Psychologe Günter Krampen; Petra Lehmann, stud. (Lehramt an Hauptschulen);
 Margarete Haag, stud. (Lehramt an Hauptschulen)
 Universität Erlangen-Nürnberg
 Fachbereich Erziehungs- und Kulturwissenschaften
 Lehrstuhl für Psychologie I
 Regensburger Str. 160
 D-8500 Nürnberg
 Tel.: 0911 / 40 27 97 (App. 187)