

Die Bewertung der Lehre durch Studierende

Ein Beispiel aus Baden-Württemberg

Hans-Dieter Daniel

Nach dem im Jahre 1998 novellierten Hochschulrahmengesetz soll die Lehre an den Hochschulen in Deutschland regelmäßig bewertet werden (§ 6). Die Studierenden sind an der Bewertung der Qualität der Lehre zu beteiligen; die Ergebnisse sollen veröffentlicht werden. Die Evaluation der Lehre erfolgt weltweit am häufigsten in Form von Hörerbefragungen (auch als Unterrichtsbeurteilung oder Veranstaltungskritik bezeichnet). Nach dem amerikanischen Hochschulforscher McKeachie (1996) war die Beurteilung von Lehrveranstaltungen durch Studierende in den letzten 30 Jahren häufiger Gegenstand der Forschung als jeder andere Aspekt des Hochschulwesens. Cashin (1995) konnte für den Zeitraum von 1971 bis 1995 mehr als 1.500 einschlägige Publikationen recherchieren. Auch in Deutschland ist in den letzten Jahren eine Fülle von Fragebögen für die Beurteilung von Lehrveranstaltungen entwickelt worden (vgl. die Dokumentation von Reissert 1992 "Evaluation der Lehre", die Beschreibung einschlägiger Projekte und Instrumente von Hage 1996, sowie die kommentierte Bibliographie "Evaluation der Hochschullehre" von Rau 1996). Musterfragebögen für Hörerbefragungen sind beispielsweise von Basler et al. (1995), Diehl (1994), Rindermann und Amelang (1994) sowie Webler, Domeyer und Schiebel (1993) publiziert worden. Gegen den Einsatz entsprechender Fragebögen wird in Deutschland eine Reihe methodenkritischer Argumente vorgetragen (vgl. z. B. Süllwold 1992).

An der Universität Mannheim stimmte der Senat bereits im Oktober 1991 auf Antrag des Allgemeinen Studierendenausschusses der Durchführung eines Modellprojekts zur Evaluierung von Lehrveranstaltungen zu, das es u. a. erlaubt, häufig geäußerte "Probleme und Gefahren" studentischer Veranstaltungskritik einer empirischen Überprüfung zu unterziehen.

1 Das öffentliche Unbehagen an den Universitäten

In einer Umfrage unter deutschen Führungspersonlichkeiten brachte das Institut für Demoskopie Allensbach das Thema Bildungssystem zur Sprache (vgl. Piel 1995). Im Vergleich zu den Fachhochschulen wurden die Universitäten verheerend beurteilt: Während über 80 % der befragten Führungskräfte der Meinung sind, dass an den Fachhochschulen alles in allem eher gute Arbeit geleistet wird, haben von den Universitäten nur 28 % der Befragten einen positiven Eindruck. Lange Studienzeiten sowie hohe Fachwechsler- und Studienabbrecherquoten

gelten vielen Außenstehenden als Indizien dafür, dass die Lehre an deutschen Universitäten eine mangelhafte Qualität besitzt (vgl. *Schröder-Gronostay; Daniel 1999*). Nicht wenige Führungskräfte teilen offenbar die Meinung des Alt-Bundeskanzlers Helmut Schmidt, dass "an der viel zu langen Studiendauer ... doch weniger die Studenten schuld (sind) als vielmehr die Hochschulorganisation" (1996).

2 Evaluationsbedarf

Dass es Mängel in der Lehre gibt, zeigen nicht zuletzt die Kommentare der Studierenden zu den von ihnen besuchten Lehrveranstaltungen. Sie beklagen, dass Pflichtveranstaltungen sich zeitlich überschneiden oder dass sie inhaltlich nicht aufeinander aufbauen. Dozenten nehmen den Begriff Vorlesung wörtlich und lesen Woche für Woche aus einem nicht selbst verfassten Buch vor oder von einem "vergilbten Manuskript" ab. Es gibt Dozenten, die in ihrem Seminar "überhaupt nichts sagen, nicht einmal in der ersten Stunde". Lehrveranstaltungen finden gar nicht statt, fallen häufig aus oder werden von den Dozenten im Laufe des Semesters regelrecht "leer gelesen". Selbst Vertreter der Standesorganisationen räumen mittlerweile ein, dass die Notwendigkeit von Evaluierungen in der Lehre genährt wird durch "Fehlritte seitens schwarzer Schafe".

Nach Auffassung des Wissenschaftsrates (1993, S. 53-55) verlangen die offenkundigen Funktionsmängel und die öffentliche Kritik an den Hochschulen nach internen und externen Verfahren der Evaluation. Hierzu gehört auch die Entwicklung von Instrumenten für die Bewertung der Lehre durch Studierende und Absolventen. Der Wissenschaftsrat empfiehlt, von Verfahren, die in anderen Ländern und an einzelnen deutschen Hochschulen üblich und allgemein akzeptiert sind, vermehrt Gebrauch zu machen: So gibt es an den meisten US-amerikanischen Universitäten schon seit Mitte der 50er Jahre Hörerbefragungen (vgl. *Ewell 1992, S. 57 f.*). An der Universität Harvard wurde Mitte der 70er Jahre die Evaluierung von Lehrveranstaltungen durch Studierende auf freiwilliger Basis eingeführt. Im ersten Jahr stimmten nur wenige Lehrende der Befragung ihrer Hörer zu. Heute werden über 80 % der Lehrveranstaltungen, die von mehr als 20 Studenten besucht werden, auf diese Weise evaluiert (*Wilkinson 1992, S. 93*). In Kanada lassen vier von fünf Universitäten regelmäßig die Lehrveranstaltungen durch die Studierenden beurteilen (*Saroyan und Donald 1994, S. 287*). In der Schweiz wurden 1993 die beiden technischen Hochschulen, die direkt dem Bund unterstellt sind, durch den ETH-Rat beauftragt, ab dem Wintersemester 1993/94 alle Lehrveranstaltungen im Turnus von drei Semestern mittels Befragung der Studierenden bezüglich ihrer didaktischen Unterrichtsqualität zu evaluieren (*Hänni 1994*). An der Deutschen Hochschule für Verwaltungswissenschaften Speyer werden von der Hörerschaft seit 1969 schriftliche Befragungen in allen Lehrveranstaltungen organisiert (*Ebling 1993*).

In der Bundesrepublik Deutschland hatte bereits im Jahre 1988 im Rahmen eines Symposiums der Westdeutschen Rektorenkonferenz eine Arbeitsgruppe unter dem Vorsitz von Friedhelm Neidhardt zum Thema "Leistungsbeurteilung und Leistungsvergleiche in der Lehre" Studenten- und Absolventenbefragungen

empfohlen (vgl. *Westdeutsche Rektorenkonferenz 1989, S. 122 f.*). Diese Empfehlung bewirkte im deutschen Hochschulsystem, das nach Heckhausen (1988, S. 34) eine "große Dickfelligkeit" und "Fähigkeit zur Selbstimmunisierung" erworben hat, zunächst nur wenig. Erst nach den Studentenprotesten im Wintersemester 1988/89 und nachdem das Nachrichtenmagazin *Der Spiegel* im Wintersemester 1989/90 die Studiensituation in den 15 meiststudierten Fächern bundesweit an 500 Fachbereichen durch Studierende beurteilen ließ, sind Befragungen der unmittelbar Betroffenen zur "Qualität der Lehre" zum Gegenstand von Förderprogrammen und Modellprojekten geworden.

3 Das Modellprojekt "Evaluation der Lehre" an der Universität Mannheim

Das Projekt verfolgte im Wesentlichen zwei Ziele: Zum einen sollten die Lehrleistungen auf einer möglichst breiten und objektiven Basis unter Einbeziehung aller Studiengänge und Lehrveranstaltungen evaluiert werden, um auf diese Weise einen Überblick über die Zufriedenheit der Studierenden mit dem Lehrangebot zu erhalten (summative Evaluation). Ferner sollten die Lehrenden zur Optimierung ihrer Lehrveranstaltungen eine differenzierte Rückmeldung seitens der Studierenden erhalten (formative Evaluation).

Im Rahmen des Modellprojekts wurden folgende Erhebungen durchgeführt:

- Hörerbefragung in den Lehrveranstaltungen „Grundzüge der Betriebswirtschaftslehre“,
- Hörerbefragung in den Lehrveranstaltungen „Statistik für Wirtschaftswissenschaftler“,
- Hörerbefragung in allen Lehrveranstaltungen der Fakultät für Sprach- und Literaturwissenschaft.

Die Befragung der Lehrveranstaltungsteilnehmer wurde um folgende Erhebungen ergänzt:

- postalische Befragung der Absolventen (vgl. *Daniel 1995, 1996*),
- postalische Befragung der Studenten im Diplom-Studiengang "Philologie mit wirtschaftswissenschaftlicher Qualifikation".

Die Antwortbereitschaft der Befragten war ausgesprochen hoch. Bei den beiden postalischen Befragungen lag die Rücklaufquote bei 80 respektive 90 %. An der Hörerbefragung in den Lehrveranstaltungen des betriebswirtschaftlichen Grundstudiums haben sich nahezu alle Studenten, die sich zu den Klausuren in "Grundzüge der Betriebswirtschaftslehre" angemeldet hatten, beteiligt. Im Rahmen der Hörerbefragung an der Fakultät für Sprach- und Literaturwissenschaft wurden je nach Fach und Abschlussart zwischen 30 und 70 % der Studierenden erreicht.

3.1 Erhebungsinstrumente

Für die Befragung der Hörer wurden studiengang- und lehrveranstaltungstyp-spezifische Fragebögen entwickelt, weil die Lehrangebote je nach Fach und Studienphase eine differenzierte Zielsetzung haben. Das Fragenprogramm bezog sich im Einzelnen auf folgende Variablenbereiche:

- individuelle Studienvoraussetzungen: z. B. Art der Hochschulreife, Abiturnote, Abschluss einer betrieblichen Lehre oder schulischen Berufsausbildung vor dem Studium;
- allgemeine Zufriedenheit mit den Studienbedingungen: z. B. technische Ausstattung der Hörsäle und der Seminarräume, Größe der Lehrveranstaltungen, Breite des Lehrangebots;
- Beurteilung der Dozenten: fachliche Kompetenz und didaktisches Geschick, Lesbarkeit von Folien, Qualität von veranstaltungsbegleitenden Materialien, Praxisrelevanz und Klausurbezug der Veranstaltungsinhalte;
- sozio-biographische Merkmale der Hörer: z. B. Studien- und Lebensalter, Geschlecht, Studiengang;
- Zufriedenheit mit dem Dozenten und mit der Lehrveranstaltung insgesamt;
- subjektiver Lernerfolg: z. B. erwartete Noten in den Abschlussklausuren;
- Prüfungserfolg in den Klausuren.

3.2 Veröffentlichung studentischer Veranstaltungskritik

Da das Datenschutzgesetz lediglich eine Veröffentlichung der Daten in aggregierter und anonymisierter Form erlaubt, bedarf die lehrveranstaltungsbezogene Veröffentlichung von Befragungsergebnissen der schriftlichen Zustimmung der Lehrenden (vgl. *Bundesdatenschutzgesetz 1995*). Nachdem die lehrveranstaltungsbezogenen Befragungsergebnisse den Lehrenden übersandt und die veranstaltungsübergreifenden Ergebnisse im Rahmen nichtöffentlicher Sitzungen den Lehrenden präsentiert und erläutert worden waren, willigten fast alle Lehrkräfte in die lehrveranstaltungsbezogene Veröffentlichung der Befragungsergebnisse ein.

3.3 Ausgewählte Ergebnisse der Hörerbefragungen

3.3.1 Zufriedenheit der Studierenden mit den Lehrveranstaltungen

Die Urteile der Studierenden über die von ihnen besuchten Lehrveranstaltungen fielen überwiegend positiv – und damit besser als erwartet – aus: Auf einer sechsstufigen Notenskala (1 = sehr zufrieden bis 6 = sehr unzufrieden) wurden die großen Vorlesungen in “Grundzüge der Betriebswirtschaftslehre” im Durchschnitt mit 3.25, die 120 vorlesungsbegleitenden Übungen und Tutorien mit 2.6 respektive 2.1 bewertet. Fasst man pro Lehrveranstaltung die Skalenwerte 1 bis 3 (zufriedene Hörer) und 4 bis 6 (unzufriedene Hörer) zusammen, dann zeigt sich, dass zwei Drittel der Vorlesungen, 91 % der vorlesungsbegleitenden Übungen und 97 % der Tutorien im positiven Bereich der Notenskala bewertet wurden.

Ebenfalls überwiegend positiv wurden die Lehrveranstaltungen "Statistik für Wirtschaftswissenschaftler" der Fakultät für Volkswirtschaftslehre und Statistik sowie die Lehrveranstaltungen der Fakultät für Sprach- und Literaturwissenschaft beurteilt. Von den zehn evaluierten Statistikkursen (mittlere Beurteilung: 2.9) erhielten acht gute bis befriedigende Beurteilungen (ein Kurs erhielt die Note 2.0, sieben Kurse Beurteilungen zwischen 2.25 und 2.9), zwei Kurse wurden mit der Note "ausreichend" beurteilt. Auch die Hörer in den 346 Lehrveranstaltungen der Fakultät für Sprach- und Literaturwissenschaft waren mit 82 % der Einführungsveranstaltungen (mittlere Beurteilung: 2.9), 96 % der Tutorien (mittlere Beurteilung: 2.0), 87 % der Vorlesungen (mittlere Beurteilung: 2.3), 94 % der wissenschaftlichen Seminare (mittlere Beurteilung: 2.4) und 89 % der sprachpraktischen Übungen (mittlere Beurteilung: 2.4) mehr oder weniger zufrieden.

3.3.2 Lehrveranstaltungsbeurteilung: Der Einfluss von potentiellen Biasvariablen

Jede Evaluation der Lehre wirft gewichtige methodische Probleme auf und ruft leicht Misstrauen, Widerstand und Kritik hervor (vgl. *Süllwold 1992*). Insbesondere die Urteilskompetenz der Studierenden wird immer wieder in Frage gestellt. Gängige Einwände sind: Die Studierenden seien zu einer objektiven Beurteilung der Lehre nicht in der Lage. Sie würden vor allem die Vorlesungen gut finden, für deren Inhalt sie sich schon vorab interessierten und deren Stoff keine allzu großen Verständnisprobleme bereite: "Ob eine Vorlesung bei den Studierenden beliebt ist oder abgelehnt wird, ob die Teilnehmer ihre Erwartungen erfüllt sehen oder nicht, das liegt nur in geringem Maße am Dozenten, an der 'Qualität der Lehre' in dieser Vorlesung, sondern es liegt in erster Linie am Stoff, der zu vermitteln ist" (*Kromrey 1993, S. 41*). Derartige Äußerungen stehen in auffälligem Kontrast zum internationalen Forschungsstand. Ganz im Sinne von Kromrey sprechen Cashin und Downey von sogenannten Biasvariablen, die sie wie folgt definieren: "the influence of variables reflecting something other than the instructor's teaching effectiveness. Such variables could be student characteristics (e. g., the students' interest in the subject matter) or course characteristics (e.g., the size of the class or the difficulty of the subject matter)" (*Cashin; Downey, 1992, S. 565*). Den Stand der Forschung zum urteilsverzerrenden Einfluss von Biasvariablen fasst Marsh wie folgt zusammen: "class-average ratings are... relatively unaffected by a variety of variables hypothesized as potential biases" (*Marsh 1984, S. 707*). Zu einer sehr ähnlichen Einschätzung kommt Cashin: "In general, student ratings tend to be statistically reliable, valid, and relatively free from bias, probably more so than any other data used for faculty evaluation" (*Cashin 1988, S. 4*). Auch die von Rindermann (1996) mit dem Heidelberger Inventar zur Lehrveranstaltungs-Evaluation (vgl. *Rindermann und Amelang 1994*) durchgeführten Hörerbefragungen haben ebenfalls keine Hinweise auf praktisch bedeutsame Biasvariablen geliefert. Eine von Greenwald (1997) durchgeführte Literaturanalyse zeigt, dass zwischen 1971 und 1995 bedeutend mehr Studien publiziert wurden, die zu dem Ergebnis kamen, dass studentische

Urteile über Lehrveranstaltungen valide sind, als Studien, die Urteilsverzerrungen (Biases) nachweisen konnten.

Um den Einfluss potentieller Biasvariablen auf die Beurteilung der Vorlesungen in "Grundzüge der Betriebswirtschaftslehre" empirisch überprüfen zu können, wurde eine multiple Regressionsanalyse mit Dummy-Variablen durchgeführt. Die Analyse zeigt, dass einige der berücksichtigten Studentenmerkmale zwar einen statistisch signifikanten Einfluss auf die Beurteilung der Vorlesungen haben. Doch weder der Studiengang noch das Studienalter (Fach- und Hochschulsesemester), die voruniversitäre Bildung (Art der Hochschulzugangsberechtigung, Abiturnote, Berufsausbildung und Berufstätigkeit vor dem Studium), das Geschlecht, das Lebensalter oder die Studienabbruchneigung üben einen praktisch bedeutsamen Einfluss auf die Beurteilung der Vorlesungen aus (Varianzanteil der Variablen jeweils < 1 %). Auch die in der Klausur erwartete Fachnote verzerrt nicht in statistisch oder praktisch bedeutsamer Weise das Urteil der Hörer über die Vorlesung (vgl. Tab. 1).

Tab. 1: Zusammenhang zwischen der erwarteten Fachnote in der Klausur BWL I bzw. BWL II und der Zufriedenheit der Hörer mit der Vorlesung insgesamt, aufgeschlüsselt nach den sechs Vorlesungen in "Grundzüge der Betriebswirtschaftslehre"

Erwartete Fachnote	„Wie zufrieden sind Sie mit dieser Vorlesung insgesamt?“					
	Vorlesung					
	A im \bar{x}	B im \bar{x}	C im \bar{x}	D im \bar{x}	E im \bar{x}	F im \bar{x}
1,0 - 2,7	2,1	2,6	2,5	2,6	3,9	4,6
2,8 - 3,3	2,2	2,6	2,6	3,4	4,2	5,2
3,4 - 3,7	2,2	2,9	2,9	2,8	4,3	4,4
$\geq 3,8$	2,2	2,8	2,8	2,2	4,6	4,7
keine Angabe	2,3	2,7	2,9	3,1	4,3	4,3
$r_{\text{erwartete Note/Zufriedenheit}}$.04	.07	.12	.05	.15	-.01

Lediglich das Interesse der Hörer am Lernstoff steht in einem statistisch (t -Wert = 14,70, $p \leq .001$) und praktisch (Varianzanteil: 17 %) bedeutsamen Zusammenhang zur Zufriedenheit der Hörer mit der Vorlesung (vgl. Tab. 2). Hörer der Vorlesungen in "Grundzüge der Betriebswirtschaftslehre" (BWL I und BWL II), die unabhängig von der Art der Vermittlung den Lernstoff der Vorlesung interessant finden, urteilen im Durchschnitt 0,7 bis 1,3 Skaleneinheiten positiver als Hörer, die den Lernstoff nicht interessant finden. Das Interesse am Lernstoff beeinflusst jedoch lediglich das Niveau, nicht jedoch die Richtung der Beurteilung: Beide Hörergruppen sind mit der Vorlesung A in der Tendenz zufriedener als mit den Vorlesungen B, C und D und mit den Vorlesungen A, B, C und D wesentlich zufriedener als mit den Vorlesungen E und F.

Tab. 2: Zufriedenheit mit der Vorlesung insgesamt, aufgeschlüsselt nach Hörern, die am Lernstoff (unabhängig von der Art der Vermittlung) interessiert respektive nicht interessiert sind

Vorlesung	Interesse am Lernstoff	
	ja Zufriedenheit mit der Vorlesung im \bar{x}	nein im \bar{x}
A	2,2	3,2
B	2,6	3,4
C	2,7	3,4
D	2,8	3,5
E	4,0	4,8
F	4,1	5,4

Anmerkungen: Prozentanteil am Lernstoff interessierter Studenten: Vorlesung A: 95 %, Vorlesung B: 87 %, Vorlesung C: 85 %, Vorlesung D: 58 %, Vorlesung E: 57 %, Vorlesung F: 63 %. Varianzaufklärung: Faktor A (Interesse) = 17 %, Faktor B (Vorlesung) = 24 %, Interaktion (AxB) = 0,3 % (sequentielle 2-faktorielle Varianzanalyse)

Das Interesse der Hörer am Lernstoff wurde auch von Marsh in den auf sehr breiter statistischer Basis durchgeführten Bias-Studien (mehr als 50.000 beurteilte Lehrveranstaltungen) als der mit Abstand bedeutsamste Einflussfaktor identifiziert: "The effect of prior subject interest on SEEQ (Students' Evaluations of Educational Quality, Anmerkung des Verfassers) scores was greater than that of any of the other background variables considered by Marsh (1980, 1983)" (Marsh und Dunkin 1992, S. 195).

Die im Rahmen des Modellprojekts "Evaluation der Lehre" an der Universität Mannheim ermittelten Ergebnisse zum Einfluss potentieller Biasvariablen auf die Beurteilung von Lehrveranstaltungen konnten von Diehl (2000) repliziert werden. Er untersuchte, inwieweit die studentische Beurteilung von Lehrveranstaltungen

staltungen im Fach Psychologie an den Universitäten in Gießen und Marburg von einer Reihe als bedeutsam erachteter Biasfaktoren abhängt. Als mögliche Biasvariablen wurden das Geschlecht der urteilenden Person (Studentinnen vs. Studenten), das Geschlecht der beurteilten Person (Dozentinnen vs. Dozenten), das Stoffgebiet der Lehrveranstaltung (methodisch vs. inhaltlich) sowie die fachliche Zusammensetzung der Teilnehmerschaft (nur Hauptfachstudenten vs. Haupt- und Nebenfachstudenten) berücksichtigt. Die Ergebnisse der Studie zeigen, dass es sich lediglich beim Faktor "Stoffgebiet der Lehrveranstaltung" um eine ernst zu nehmende Biasvariable handelt.

Angesichts der vorliegenden empirischen Evidenz scheinen Zweifel an der Fairness studentischer Urteile weitgehend unbegründet zu sein. Sofern von einzelnen Personen- oder Veranstaltungsmerkmalen (z. B. Interesse am Lernstoff, Größe der Lehrveranstaltung) ein urteilsverzerrender Effekt ausgeht, kann er – wie oben gezeigt wurde – auf statistischem Wege kontrolliert werden. Kritischer als die Fairness ist die Reliabilität studentischer Urteile einzuschätzen. Die Ergebnisse der an der Universität Mannheim durchgeführten Hörerbefragungen haben deutlich gemacht, dass verschiedene Hörergruppen die gleiche Lehrveranstaltung (in Form von Parallelkursen angeboten) recht unterschiedlich einschätzen können (vgl. Tab. 3). Dozent A beispielsweise hat fünf parallele Übungen zu einer der sechs Vorlesungen in "Grundzüge der Betriebswirtschaftslehre" veranstaltet. Auf einer sechsstufigen Notenskala wurde die Übung 1 des Dozenten mit zwölf Hörern im Durchschnitt mit 1.6 bewertet. Die Übung 5 mit 54 Hörern wurde demgegenüber im Durchschnitt mit 2.5 beurteilt. Etwa die Hälfte aller Dozenten wurde in ihren Parallelveranstaltungen von den jeweiligen Hörergruppen statistisch signifikant unterschiedlich beurteilt (d. h. auf der sechsstufigen Notenskala variierte die mittlere Zufriedenheit mit der Lehrveranstaltung um eine ganze Note). Ähnliche Erfahrungen berichtet Greenwald (1997, S. 1182): "My interest in student ratings had a sudden onset. In 1989, I received the highest student rating evaluations I had ever received at University of Washington, for teaching an undergraduate honors seminar. The sudden interest came, not then, but a year later, when I received my lowest ever evaluations. The two ratings were separated by eight deciles according to the university's norms – about 2.5 standard deviations apart. But these two ratings were for the same course, taught in the same fashion, with a syllabus that was only slightly changed". Die geringe Reliabilität der studentischen Zufriedenheitsurteile hat zur Folge, dass von einem "Dozenten-Ranking" abgesehen werden muss. In unserem Beispiel sind die Dozenten A bis J in etwa gleich gut beurteilt worden. Mit den Lehrveranstaltungen der Dozenten K und L waren die Studierenden allerdings deutlich weniger zufrieden.

Tab. 3: Mittlere Zufriedenheit verschiedener Hörergruppen mit identischen vorlesungsbegleitenden Übungen (Parallelveranstaltungen)

Dozent		Übung (Parallelveranstaltung)								
		1	2	3	4	5	6	7	8	
A	im \bar{x}	1,6	1,9	2,0	2,3	2,5				
	N	12	53	49	68	54				
B	im \bar{x}	2,0	2,1	2,2	2,4					
	N	42	32	22	37					
C	im \bar{x}	2,0	2,1	2,2	2,4					
	N	29	41	74	99					
D	im \bar{x}	2,3	2,4	2,5						
	N	108	50	29						
E	im \bar{x}	2,3	2,3	2,5	2,5					
	N	125	78	49	106					
F	im \bar{x}	1,8	2,1	2,3	2,4	2,5	2,5	2,6	2,7	
	N	17	42	65	64	111	60	120	25	
G	im \bar{x}	2,0	2,2	2,4	2,4	2,5	2,7	2,7	2,9	
	N	19	45	58	55	61	102	121	25	
H	im \bar{x}	2,4	2,5	2,9	2,9	3,2				
	N	48	34	18	71	47				
I	im \bar{x}	2,7	2,8	2,8	3,0					
	N	66	79	59	72					
J	im \bar{x}	2,4	3,2	3,3	3,3					
	N	24	53	62	67					
K	im \bar{x}	3,2	3,3	3,6	3,7					
	N	39	70	56	24					
L	im \bar{x}	3,8	4,5	5,0						
	N	4	2	3						

Anmerkung: Die Frage lautete: „Wie zufrieden sind Sie mit dieser Übung insgesamt?“ (Verankerung der Antwortskala: 1 = sehr zufrieden, 6 = sehr unzufrieden). Die Parallelveranstaltungen der Dozenten A, F, G, H und J wurden von den jeweiligen Hörergruppen statistisch signifikant ($p > .05$), die der Dozenten C und K in der Tendenz ($p > .10$) unterschiedlich beurteilt.

Süllwold (1992) hat darauf hingewiesen, dass unter methodischen Gesichtspunkten ein Vergleich von Lehrveranstaltungen immer dann problematisch ist, wenn die Lehrveranstaltungen von verschiedenen Studentengruppen beurteilt werden. Die günstigsten Voraussetzungen für einen Vergleich wären gegeben, wenn alle zu vergleichenden Lehrveranstaltungen von der gleichen Studentengruppe beurteilt würden. An den Universitäten ist dieser Untersuchungsplan allerdings schwer zu realisieren, weil – zumindest in den wenig regulierten Studiengängen

– nur eine relativ kleine Gruppe von Studenten dieselbe Konfiguration von Lehrveranstaltungen besucht. Das von Süllwold (1992) beschriebene Untersuchungsdesign würde es auch erlauben, die Übereinstimmungsreliabilität der studentischen Urteile zu überprüfen.

Die Übereinstimmungsreliabilität (ausgedrückt im Intraclass-Korrelations-Koeffizienten) konnte anhand von Daten überprüft werden, die dem Autor dankenswerterweise von der Bankakademie in Frankfurt am Main zur Verfügung gestellt wurden. An der Bankakademie besucht eine Jahrgangskohorte (Kurs) jeweils dieselben Lehrveranstaltungen in den fünf Fächern Bankbetriebslehre, Betriebswirtschaftslehre, Volkswirtschaftslehre, Betriebliches Rechnungswesen und Personalwesen. Die Lehrveranstaltungen werden von den Kursteilnehmern hinsichtlich Stoffgliederung, verständlicher Darstellung der Vorlesungsinhalte, Verdeutlichung des Stoffes durch Fälle/Aufgaben, Medieneinsatz (Projektor, Tafel), Möglichkeiten zur aktiven Mitarbeit und Engagement des Dozenten vergleichend beurteilt. Tabelle 4 zeigt, dass die Intraclass-Korrelations-Koeffizienten – je nach Beurteilungsdimension und Kurs – zwischen .00 und .46 liegen. Intraclass-Korrelations-Koeffizienten in dieser Höhe drücken eine relativ geringe Urteilskonkordanz aus. Aus statistischer Sicht muss die Übereinstimmungsreliabilität der studentischen Urteile deshalb als wenig befriedigend bezeichnet werden. Dies gilt im übrigen auch für die Übereinstimmung gutachterlicher Stellungnahmen von Wissenschaftlern für Fachzeitschriften (vgl. Daniel 1993).

Tab. 4: Übereinstimmungsreliabilität studentischer Urteile über Lehrveranstaltungen: Beurteilung von fünf Lehrveranstaltungen (BBL, BWL, VWL, BR, PW) an einer Bankakademie durch jeweils identische Hörergruppen (Intraclass-Korrelations-Koeffizienten)

Beurteilungs-Dimension	Kurs A (N = 12 Hörer)	Kurs B (N = 24 Hörer)
Stoffgliederung	.23	.19
Verständliche Darstellung der Vorlesungsinhalte	.46	.00
Verdeutlichung des Stoffes durch Fälle/Aufgaben	.35	.07
Medieneinsatz (Projektor, Tafel)	.35	.14
Möglichkeiten zur aktiven Mitarbeit	.35	.45
Engagement des Dozenten	.45	.37

3.3.3 Korrelate der individuellen Zufriedenheit

Die Zufriedenheit der Studierenden mit der Vorlesung hängt sehr stark davon ab, ob die Dozenten in ihrer Lehrveranstaltung den Eindruck erwecken, gut vorbereitet zu sein und den Gegenstand der Vorlesung souverän zu beherrschen. Neben diesen fachlichen Aspekten hat jedoch auch die Präsentation des Lernstoffs ("Der Dozent versteht es, die Hörer für den Stoff der Vorlesung zu interessieren", "Der Dozent gestaltet die Vorlesung lebendig und engagiert", "Der Dozent gliedert den Stoff der Vorlesung übersichtlich", "Komplizierte Sachverhalte werden verständlich erläutert", "Der Dozent fasst den behandelten Stoff gut zusammen") und die soziale Kompetenz des Lehrenden ("Der Dozent sorgt für eine angenehme Atmosphäre in der Vorlesung") einen großen Einfluss auf die Zufriedenheit der Hörer (Beispiele für fachliche, didaktische und interpersonale Defizite finden sich im Übrigen in großer Zahl in den ausführlichen Fragebogenkommentaren der Studierenden zu den einzelnen Lehrveranstaltungen). Dies deutet darauf hin, dass für die zufriedenstellende Vermittlung des Lehrstoffs neben dem Fachwissen auch die didaktischen und sozialen Fertigkeiten des Dozenten von Bedeutung sind (vgl. hierzu auch Preißer 1991, S. 423 ff.). Anders ausgedrückt: Die Studierenden erwarten von ihrem Dozenten mehr als nur das wörtliche Vorlesen aus einem Lehrbuch ("Lesen können wir auch zu Hause"). Die Studierenden in den Lehrveranstaltungen des betriebswirtschaftlichen Grundstudiums erwarten vor allem, dass sie in der Vorlesung, die sie besuchen, etwas lernen. Die Antworten der Hörer auf die Frage "Wie viel haben Sie Ihrer Meinung nach in dieser Vorlesung gelernt?" sind in allen Vorlesungen mit Abstand der beste Prädiktor für die individuelle Zufriedenheit mit der Lehrveranstaltung ($r = .77$, $N = 2295$). In allen Vorlesungen sind die Studierenden um so zufriedener, je mehr sie in der Vorlesung meinen, gelernt zu haben.

Sortiert man die sechs Vorlesungen nach dem Prozentanteil der Hörer, die der Meinung sind, in der Vorlesung mehr oder weniger viel gelernt zu haben (Skalenwerte 1-3), dann erhält man exakt die gleiche Rangfolge der Vorlesungen wie nach der gemittelten individuellen Zufriedenheit der Hörer mit der Vorlesung insgesamt (vgl. Tab. 6). Während in den Vorlesungen A bis D zwischen 90 und 68 % der Hörer meinen, mehr oder weniger viel gelernt zu haben, sind es in den Vorlesungen E und F nur 37 respektive 29 % der Hörer.

Tab. 5: Zusammenhang zwischen der Menge des in der Vorlesung subjektiv Gelernten und der Zufriedenheit der Hörer mit der Vorlesung insgesamt, aufgeschlüsselt nach den sechs Vorlesungen in "Grundzüge der Betriebswirtschaftslehre"

„Wie zufrieden sind Sie mit dieser Vorlesung insgesamt?“						
„Wieviele haben Sie Ihrer Meinung nach in dieser Vorlesung gelernt?“	Vorlesung					
	A im ø	B im ø	C im ø	D im ø	E im ø	F im ø
sehr viel	1,5	1,7	1,8	1,9	1,5	1,6
2	2,0	2,2	2,3	2,4	2,5	2,1
3	2,4	3,0	2,8	2,9	3,6	2,7
4	3,0	3,8	3,2	3,6	4,4	4,1
5	3,5	4,4	4,0	4,5	5,1	5,1
sehr wenig	4,8	-	4,2	5,3	5,9	5,8
r	.53	.57	.51	.70	.77	.88

Anmerkung: Varianzaufklärung: Faktor A (Menge des in der Vorlesung subjektiv Gelernten) = 60 %, Faktor B (Vorlesung) = 5 %, Interaktion (AxB) = 2 %

Tab. 6: Prozentanteil der Hörer in den sechs Vorlesungen in "Grundzüge der Betriebswirtschaftslehre", die ihrer Meinung nach in der Vorlesung viel (Skalenwerte 1-3) respektive wenig (Skalenwerte 4-6) gelernt haben

„Wie viel haben Sie Ihrer Meinung nach in dieser Vorlesung gelernt?“		
Vorlesung	Viel	Wenig
A	90	10
B	88	12
C	77	23
D	68	32
E	37	63
F	29	71

Anmerkung: Chi²-Wert = 520, FG = 5, p ≤ .001.

Auch die Antworten der Hörer auf die Frage "Wie gut fühlen Sie sich durch die Vorlesung auf die Klausur in BWL I bzw. BWL II vorbereitet?" korrelieren hoch mit der individuellen Zufriedenheit: Je besser sich die Studierenden durch die Vorlesung auf die Klausur vorbereitet fühlen, desto zufriedener sind sie mit der Lehrveranstaltung ($r = .61$). [Demgegenüber ist die erwartete Klausurnote nahezu unkorreliert mit der Hörerzufriedenheit (vgl. Tab. 1). Dies kann als Beleg für die Fairness der studentischen Lehrveranstaltungskritik gewertet werden.] Die Klausuren selbst sind Gegenstand heftiger studentischer Kritik. Kritisiert werden vor allem die hohen Durchfallquoten, die rechtsschiefe Notenverteilung (vgl. Abb. 1 und 2) und die – im Vergleich zu den erzielten Abiturnoten (vgl. Abb. 3) – mit 3.9 und 4.0 relativ schlechten Durchschnittsnoten.

Abb. 1: Notenverteilung in der Klausur BWL I (n = 763)

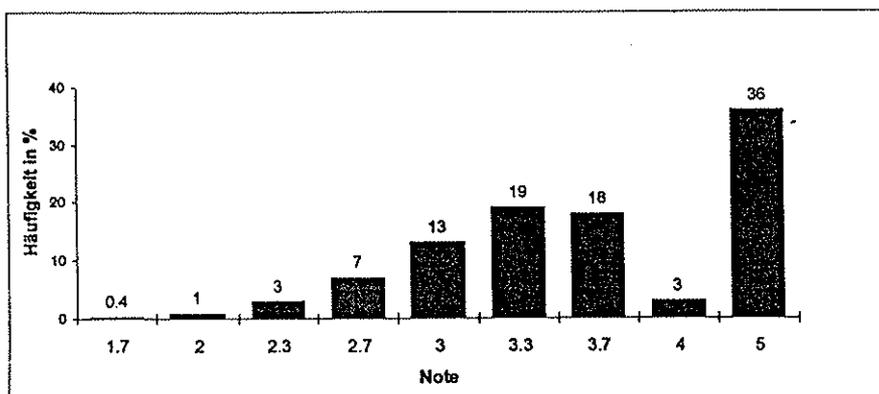


Abb. 2: Notenverteilung in der Klausur BWL II (n = 559)

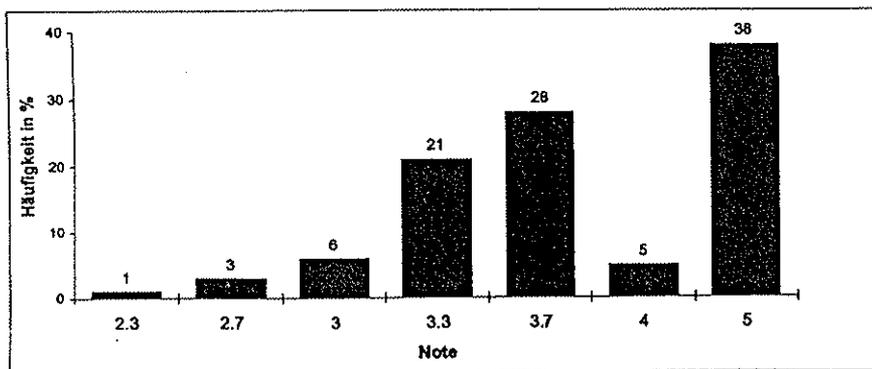
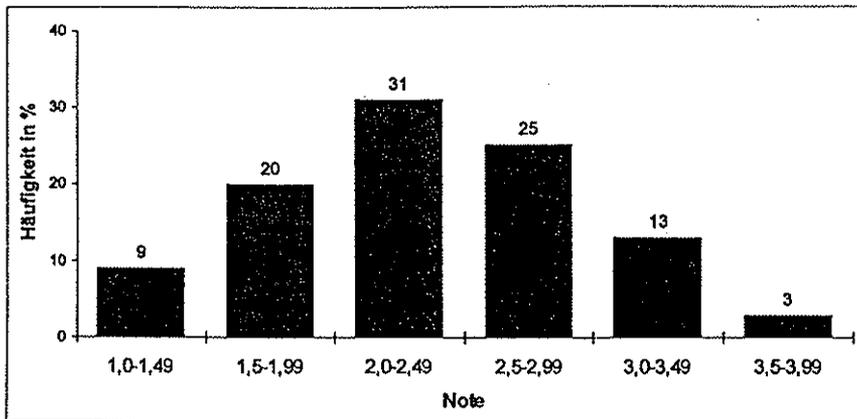


Abb. 3: Verteilung der Abiturnote der befragten Studierenden (n = 1175)



Die Auswertung der Klausurergebnisse zeigt jedoch, dass die studentische Kritik an der Notengebung nur zum Teil berechtigt ist. Tabelle 7 zeigt, dass der Prüfungserfolg der Studierenden sehr gut aufgrund der Abiturnote vorhergesagt werden kann (vgl. hierzu Schuler 1989 sowie Baron-Boldt; Funke; Schuler 1989). Je besser die Abiturnote der Studierenden, um so besser das in den Klausuren erzielte Ergebnis (die Korrelation zwischen der Abiturnote und dem Klausurergebnis ist in BWL II erwartungsgemäß niedriger als in BWL I, weil durch die selektierende Wirkung der Klausur in BWL I die leistungsschwächsten Studierenden an der Klausur in BWL II nicht mehr teilnehmen). Damit die Studierenden durch die Notengebung nicht demotiviert werden, könnte allerdings die Notenskala besser ausgenutzt, das Niveau der Benotung geändert oder den Studierenden die Notenverteilung sowie die Rangplatzziffer bezogen auf die Prüfungskohorte mitgeteilt werden [um die von den Studierenden beklagte rechtschiefe Notenverteilung ("keine Einser") zu beseitigen, hat die Fakultät für Betriebswirtschaftslehre ab dem Wintersemester 1994/95 den Notenschlüssel geändert] (vgl. Daniel 1995, S. 34).

Tab. 7: Zusammenhang zwischen der Abiturnote und der Note in den Klausuren in BWL I und BWL II

Abiturnote	Durchschnittsnote in ...	
	BWL I	BWL II
1,0-1,4	2,9	3,4
1,5-1,9	3,3	3,7
2,0-2,4	3,7	3,9
2,5-2,9	4,0	4,1
3,0-3,4	4,2	4,2
3,5-4,0	4,7	4,4
r	.49	.26

3.3.4 Perzipierte Studienbarrieren

Die Zufriedenheit der Studierenden mit den besuchten Lehrveranstaltungen hängt nach den oben referierten Befunden in hohem Maße von der fachlichen, didaktischen und sozialen Kompetenz der Lehrenden ab. Während die fachliche Qualifikation der Dozenten im Allgemeinen sehr gelobt wird, monieren die Studierenden von Fall zu Fall die unzureichende didaktische Aufbereitung des Lehrstoffs und Defizite in der sozialen Kompetenz einzelner Dozenten. Als ausschlaggebender Grund für den Besuch von Lehrveranstaltungen wird die Fachkompetenz allerdings weit häufiger genannt als die didaktische oder soziale Kompetenz der Lehrenden (vgl. hierzu auch *Schweer; Rosemann 1995, S. 190 f.*). Da in allen untersuchten Fakultäten die Hörer mit mehr als 80 % der Lehrveranstaltungen zufrieden sind, verwundert es nicht, dass "fachlich besser qualifizierte Dozenten" und "mehr Engagement der Dozenten in den Seminaren" nach Meinung der Studierenden am wenigsten dringlich erscheinen, um die Studiensituation zu verbessern (die Wunschliste umfasste insgesamt 39 Antwortvorgaben). Moniert wurden von den Studierenden des betriebswirtschaftlichen Grundstudiums vor allem technische Probleme (Akustik und Videoprojektion) in den Lehrveranstaltungsräumen, die Prüfungsregelungen und die Notengebung in den Klausuren, die Studieninhalte sowie die großen Unterschiede im Vorwissen der heterogenen Hörschaft.

Die Hörer der Lehrveranstaltungen haben eine Reihe von Vorschlägen zur Verbesserung von Lehre und Studium unterbreitet. Die studentischen Verbesserungsvorschläge beziehen sich auf folgende Bereiche:

1. Prüfungsregelungen und Notengebung

Folgende Verbesserungsvorschläge wurden unterbreitet:

- Häufigere Leistungskontrollen, beispielsweise in Form von abgabepflichtigen Hausaufgabenblättern und Übungsklausuren;
- für jedes Fach eine separate Prüfung statt einer Blockprüfung;
- Wiederholungsklausur am Ende der vorlesungsfreien Zeit (und nicht erst im folgenden Semester);
- Ausschöpfung der gesamten Notenskala;
- Verringerung der Durchfallquoten.

2. Praxisbezug des Studiums

Folgende Verbesserungsvorschläge wurden unterbreitet:

- häufiger Gastdozenten aus der Wirtschaft einladen;
- häufiger Betriebsbesichtigungen durchführen;
- häufiger Fallstudien durchführen;
- Einführung eines Praxissemesters;
- Kooperation mit Mentorenfirmen.

3. Studieninhalte

Die Verbesserungsvorschläge beziehen sich vor allem auf das Fach “Grundzüge der Volkswirtschaftslehre”:

- VWL-Inhalte im BWL-Studium verringern;
- Verzicht auf übertriebene Mathematik in VWL;
- Verzicht auf Multiple-Choice-Klausuren in VWL.

4. Heterogene Hörerschaft

Folgende Verbesserungsvorschläge wurden unterbreitet:

- Die Lehrenden des Faches Betriebswirtschaftslehre sollten stärker berücksichtigen, dass etwa die Hälfte der Hörer ein anderes Hauptfach studiert und Kenntnisse in Mathematik, Recht oder VWL nicht in gleicher Weise wie bei Studierenden des Diplom-Studienganges Betriebswirtschaftslehre vorausgesetzt werden können. Angeregt werden vorlesungsbegleitende Übungen und Tutorien für homogene Teilgruppen der Hörerschaft.

5. Zulassungsbeschränkungen

Von einem Teil der Studierenden wird gefordert, den Zugang zum Studium der Betriebswirtschaftslehre stärker zu reglementieren. Für die Studierenden der

Fakultät für Sprach- und Literaturwissenschaft stehen dagegen folgende Probleme und Verbesserungsvorschläge im Vordergrund:

- Kleines Latein: Jeder zweite Student, dessen Prüfungsordnung Lateinkenntnisse vorschreibt, muss das Latein während des Studiums nachholen (*zur Diskussion um Latein als Schulsprache vgl. Haag 1995*). Nach Einschätzung der Studierenden verlängert dies die Studienzzeit um durchschnittlich 1.5 Semester. Nach Meinung der Mehrheit der Studierenden sollte in den Magisterstudiengängen Latein durch eine zweite Fremdsprache substituierbar sein. Sofern Lateinkenntnisse unverzichtbar sind, sollte deren Umfang auf die speziellen Studierfordernisse (etwa im Fach Sprachgeschichte) beschränkt werden.
- Zulassung zu sprachpraktischen Übungen: Die Zulassung zu sprachpraktischen Übungen sollte nach dem Studienstand und nicht aufgrund eines Losverfahrens ("Roulette") erfolgen.
- Studium einer Nicht-Schulsprache: Studenten, die in einem Nicht-Schulfach eingeschrieben sind, wünschen sich ein zusätzliches Angebot von propädeutischen Fremdsprachenkursen in der vorlesungsfreien Zeit.
- Zeitliche Überschneidung von Pflichtveranstaltungen: In den Magister- und Lehramtsstudiengängen klagen viele Studenten über zeitliche Überschneidungen von Pflichtveranstaltungen (insbesondere, wenn mehrere Fakultäten für einen Studiengang zuständig sind). Durch den Einsatz computerunterstützter Planungs- und Entscheidungshilfen (*vgl. Ferland; Fleurent 1994*) könnten zeitliche Überschneidungen von Pflichtveranstaltungen vermieden werden.
- Struktur des Lehrangebots: Die Studierenden wünschen sich eine andere Struktur des Lehrangebots: weniger Seminare (in denen überwiegend Spezialwissen vermittelt wird) und mehr Vorlesungen (zur Vermittlung von Überblicks- und Zusammenhangswissen).
- Studium in der Fremdsprache: In den fremdsprachlichen Philologien fordern die Studierenden eine bessere Vorbereitung auf die mündliche Abschlussprüfung (indem beispielsweise mehr Vorlesungen in der Fremdsprache angeboten werden und in den Seminaren häufiger in der Fremdsprache referiert und diskutiert wird).

Darüber hinaus fordert ein großer Teil der Studierenden eine Änderung des Prüfungssystems (Sukzessiv- statt Blockprüfung), eine bessere Ausstattung der Bibliotheken, mehr und besser dotierte Auslandsstipendien, höhere BAföG-Sätze und Stipendien (70 % der befragten Studierenden können sich nicht ausreichend auf das Studium konzentrieren, weil sie zur Finanzierung ihres Lebensunterhalts während des Semesters erwerbstätig sein müssen) sowie eine bessere Orientierung über berufliche Einsatzfelder insbesondere für Geisteswissenschaftler.

Einige der o. g. Vorschläge sind auch von den Fachkommissionen der Hochschulrektorenkonferenz unterbreitet worden, welche die fachspezifischen Gründe für die langen Studienzeiten analysiert haben (vgl. *Hochschulrektorenkonferenz 1994*): So wird für den Magisterstudiengang Anglistik beispielsweise vorgeschlagen, dass Latein durch umfassende Kenntnisse in Französisch substituierbar sein sollte (S. 13). Im Studiengang Betriebswirtschaftslehre spricht sich die Fachkommission für zeitnahe Wiederholungsprüfungen aus (S. 44).

4 Fazit: Qualitätssicherung durch Evaluierung der Lehre

In den bisher in Deutschland durchgeführten Studien kommen die Autoren übereinstimmend zu dem Ergebnis, dass die Studierenden mit der großen Mehrzahl der von ihnen besuchten Lehrveranstaltungen im Großen und Ganzen zufrieden sind. Zu einem sehr ähnlichen Ergebnis kommen die zahlreichen anglo-amerikanischen Studien: "student ratings typically tend to be favorable" (*Educational Testing Service o. J.*). Trotz der überwiegend positiven Beurteilung von Lehrveranstaltungen besteht jedoch kein Grund zu übertriebener Selbstzufriedenheit. Zu Recht kann eingewendet werden, dass Untersuchungen über die Zufriedenheit von Lehrveranstaltungsteilnehmern nur bedingt aussagekräftig sind:

- Bei Zufriedenheitsbefragungen besteht im Allgemeinen eine starke Tendenz zu sehr positiven Antworten (vgl. *Aust 1994*).
- Zufriedenheit ist in hohem Maße vom individuellen Anspruchsniveau abhängig. Aus diesem Grunde kann Zufriedenheit auch dann geäußert werden, wenn die Bedingungen objektiv schlecht sind (der Dozent beispielsweise im Seminar "überhaupt nichts sagt" oder von einem "vergilbten Manuskript" abliest).
- Ferner schließt eine positive Antwort auf die Frage "Wie zufrieden sind Sie mit der Lehrveranstaltung insgesamt?" erhebliche Detailkritik (etwa an der mangelnden Transparenz der Prüfungsanforderungen) nicht aus.

Auch ist der Einwand nicht von der Hand zu weisen, dass Hörerbefragungen wahrscheinlich ein zu positives Bild der Lehr- und Studiensituation zeichnen. Studierende, die während des Semesters den Besuch von Lehrveranstaltungen abbrechen, sowie Studierende, die – aus welchen Gründen auch immer – überhaupt keine Lehrveranstaltungen besuchen, werden durch Hörerbefragungen nicht erreicht. Die Ergebnisse der an der Fakultät für Sprach- und Literaturwissenschaft sowie an der Fakultät für Sozialwissenschaften an der Universität Mannheim durchgeführten flächendeckenden Hörerbefragungen legen den Schluss nahe, dass die Hälfte aller immatrikulierten Studenten keine Lehrveranstaltungen besucht (vgl. *Esser 1995; Bauer 1997*).

Trotz der genannten Einwände stellen Befragungen von Studierenden für jede Hochschule unverzichtbare Evaluations- und Qualitätssicherungsinstrumente dar, weil sie öffentlich sichtbar machen, wie die Lehre von den unmittelbar Betroffenen angenommen wird. Darüber hinaus tragen sie dazu bei, Studienbarrie-

ren und damit die Gründe für überlange Studienzeiten, hohe Fachwechsler- und Studienabbrecherquoten zu identifizieren. Methodische Bedenken sollten die Hochschulen nicht länger davon abhalten, Schwachstellen im Bereich von Lehre und Studium wissenschaftlich zu analysieren und allfällige Reformen einzuleiten.

Literaturverzeichnis

Aust, B. (1994): Zufriedene Patienten? Eine kritische Diskussion von Zufriedenheitsuntersuchungen in der gesundheitlichen Versorgung. Berlin: Wissenschaftszentrum Berlin für Sozialforschung 1994

Baron-Boldt, J.; Funke, U.; Schuler, H. (1989): "Prognostische Validität von Schulnoten. Eine Metaanalyse der Prognose des Studien- und Ausbildungserfolgs". In: Jäger, Reinhold S.; Horn, R.; Ingenkamp, K. (Hg.): Tests und Trends 7. Jahrbuch der Pädagogischen Diagnostik. Weinheim: Beltz 1989, S. 11-39

Basler, H.-D.; Bolm, G.; Dickescheid, T.; Herda, C. (1995): Marburger Fragebogen zur Akzeptanz der Lehre. Diagnostica, 41. Jg., 1995, H. 1, S. 62-79

Bauer, H. H.: The New Generations: Students Who Don't Study. Internet-URL: <http://www.bus.lsu.edu/accounting/faculty/lcrumbley/study.htm>

Bundesdatenschutzgesetz: 4., neubearbeitete Auflage. München: C. H. Beck 1995

Cashin, W. E.: Student Ratings of Teaching: The Research Revisited. Kansas State University, Centre for Faculty Evaluation & Development (IDEA Paper No. 32)

Cashin, W. E.; Downey, R. G. (1992): "Using Global Student Rating Items for Summative Evaluation". In: Journal of Educational Psychology, 84. Jg., 1992, H. 4, S. 563-572

Daniel, C. (1995): "Was bringen die neuen Prüfungsregelungen in BWL I und II?". In: transmitter – Studentinnen- und Studentenzeitung der Universität Mannheim, 13. Jg., 1995, H. 51, S. 34

Daniel, H.-D. (1993): Guardians of Science – Fairness and Reliability of Peer Review. Weinheim: VCH 1993

Daniel, H.-D. (1996): "Korrelate der Fachstudiendauer von Betriebswirten: Ergebnisse einer Absolventenbefragung an der Universität Mannheim". In: Albach, H.; Brockhoff, K. (Hg.), Betriebswirtschaftslehre und der Standort Deutschland. Zeitschrift für Betriebswirtschaft (Ergänzungsheft 1/96), 65. Jg., 1996, H. 1, S. 95-115

Daniel, H.-D. (1995): "Der Berufseinstieg von Betriebswirten". In: Personal- Zeitschrift für Human Resource Management, 47. Jg., 1995, H. 10, S. 492-499

Diehl, J. M. (1994): Fragebogen zur studentischen Evaluation von Hochschulveranstaltungen. Manual zum VBVOR und VBREF. Fachbereich 06 Psychologie (Abteilung Methodik). Universität Gießen 1994

- Diehl, J. M. (2000):* "Studentische Lehrevaluation in den Sozialwissenschaften: Fragebögen, Normen, Probleme". In: Keiner, E. (Hg.): Evaluation (in) der Erziehungswissenschaft. Weinheim: Deutscher Studien Verlag 2000 (im Druck)
- Ebling, W. (1993):* "Schriftliche Hörerbefragung an der Hochschule für Verwaltungswissenschaften". In: Barz, Andreas (Hrsg. im Auftrag des Ministeriums für Wissenschaft und Weiterbildung Rheinland-Pfalz): Qualität der Lehre – Verfahren der Evaluation an Hochschulen in Rheinland-Pfalz. Kaiserslautern 1993, S. 31-32
- Educational Testing Service: Guidelines for the Use of Results of the Student Instructional Report.* Princeton: ETS/Higher Education Assessment o. J.
- Esser, H. (1995):* "Lehrevaluation – No Shows, Karteileichen, Schleifendreher". In: Deutsche Universitäts-Zeitung, 1995, H. 18, S. 22-25
- Ewell, Peter T. (1992):* "Lehrevaluation in den USA – ein Wegweiser durch die Vielfalt neuer Assessment-Ansätze". In: Holtkamp, R. und Schnitzer, K. (Hg.): Evaluation des Lehrens und Lernens: Ansätze, Methoden, Instrumente – Evaluationspraxis in den USA, Großbritannien und den Niederlanden. Hannover: HIS GmbH 1992, S. 51-64
- Ferland, J. A.; Fleurent, Ch. (1994):* "SAPHIR: A Decision Support System for Course Scheduling". In: Interfaces, 24. Jg., 1994, S. 105-115
- Greenwald, A. G. (1997):* "Validity Concerns and Usefulness of Student Ratings of Instruction". In: American Psychologist, 52. Jg., 1997, H. 11, S. 1182-1186
- Haag, L. (1995):* "Auswirkungen von Lateinunterricht". In: Psychologie in Erziehung und Unterricht, 42. Jg., 1995, S. 245-254
- Hänni, H. (1994):* Unterrichtsbeurteilung durch Studierende an der ETH Zürich. Zürich: ETH Didaktikzentrum 1994
- Hage, N. (1996):* Lehrevaluation und studentische Veranstaltungskritik – Projekte, Instrumente und Grundlagen. Bonn: Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie 1996
- Heckhausen, H. (1988):* "Wozu ortsvergleichende Produktivitäts-Ranglisten von Fachbereichen und wie deshalb vorgegangen werden sollte". In: Daniel, H.-D.; Fisch, R. (Hg.): Evaluation von Forschung. Konstanz: Universitätsverlag Konstanz 1988, S. 33-39
- Hochschulrektorenkonferenz (1994):* Studienstrukturreform in Anglistik und Betriebswirtschaftslehre. Bonn 1994 (Dokumente zur Hochschulreform; 92)
- Kromrey, H. (1993):* "Studentische Vorlesungskritik. Empirische Daten und Konsequenzen für die Lehre". In: Soziologie, 1993, H. 1, S. 39-56
- Marsh, H. W. (1980):* "The Influence of Student, Course and Instructor Characteristics on Evaluations of University Teaching". In: American Educational Research Journal, 17. Jg., 1980, S. 219-237
- Marsh, H. W. (1983):* "Multidimensional Ratings of Teaching Effectiveness by Students from Different Academic Settings and Their Relation to Student/Course/Instructor Characteristics". In: Journal of Educational Psychology, 75. Jg., 1983, S. 150-166
- Marsh, H. W. (1984):* "Students' Evaluations of University Teaching. Dimensionality, Reliability, Validity, Potential Biases, and Utility". In: Journal of Educational Psychology, 76. Jg., 1984, H. 5, S. 707-754

- Marsh, H. W.; Dunkin, M. J. (1992): "Students' Evaluations of University Teaching. A Multidimensional Perspective". In: SMART, John C. (Hg.): Higher Education. Handbook of Theory and Research. New York: Agathon 1992, Volume VIII, S. 143-233*
- McKeachie, W. J.: Student Ratings of Teaching. University of Michigan 1996. Internet URL: <http://www.acls.org/op33.htm#McKeachie>*
- Piel, E. (1995): "Über das Unbehagen an unserem Bildungssystem: Ergebnisse einer Umfrage des Instituts für Demoskopie Allensbach". In: Forschung & Lehre (MittHV), 41. Jg., 1995, S. 5-7*
- Preißer, R. (1991): "Die Qualität der Lehre: Ergebnisse einer Befragung an der Technischen Universität Berlin". In: Beiträge zur Hochschulforschung, 1991, H. 4, S. 401-429*
- Rau, E. (1996): Evaluation der Hochschullehre – Eine kommentierte Bibliographie. Frankfurt am Main: Peter Lang 1996*
- Reissert, R. (1992): Dokumentation: Evaluation der Lehre – Aktuelle Aktivitäten an deutschen Hochschulen (Teil 1 + 2). Hannover: Hochschul-Informationssystem GmbH 1992*
- Rindermann, H. (1996): Untersuchungen zur Brauchbarkeit studentischer Lehrevaluationen. Landau: Verlag Empirische Pädagogik 1996*
- Rindermann, H.; Amelang, M. (1994): Das Heidelberger Inventar zur Lehrveranstaltungsevaluation (HILVE). Heidelberg: Asanger 1994*
- Saroyan, A.; Donald, J. G. (1994): "Evaluating Teaching Performance in Canadian Universities". In: Das Hochschulwesen, 1994, H. 6, S. 282-291*
- Schmidt, H. (1996): "Rezepte gegen die deutsche Krankheit". In: Die Zeit, Nr. 3 vom 12. Januar 1996, S. 3*
- Schröder-Gronostay, M.; Daniel, H.-D. (1999): Studierenerfolg und Studienabbruch – Beiträge aus Forschung und Praxis. Neuwied: Luchterhand 1999*
- Schuler, H. (1989): "Gute Abiturnoten, aber schlechter Studierenerfolg? Schulnoten als Prognosegrundlagen für das Hochschulstudium". In: Mitteilungen des Hochschulverbandes, 37. Jg., 1989, H. 6, S. 321-322*
- Schweer, M. K. W.; Rosemann, B. (1995): "Qualität der Lehre: Bedingungsvariablen des studentischen Urteils". In: Zeitschrift für Pädagogische Psychologie, 9. Jg., 1995, S. 189-196*
- Süllwold, F. (1992): "Universitäre Lehre. Welche Realität wird bei der Beurteilung von Hochschullehrern durch Studierende erfaßt?". In: Mitteilungen des Hochschulverbandes, 40. Jg., 1992, H. 1, S. 34-35*
- Webler, W.-D.; Domeyer, V.; Schiebel, B. (1993): Lehrberichte – Empirische Grundlagen, Indikatorenauswahl und Empfehlungen zur Darstellung der Situation der Lehre in Lehrberichten. Bonn: Bundesministerium für Bildung und Wissenschaft 1993 (Schriftenreihe Studien zu Bildung und Wissenschaft 107)*
- Westdeutsche Rektorenkonferenz (1989): Leistungsbeurteilung und Leistungsvergleich im Hochschulbereich. Bonn 1989 (Dokumente zur Hochschulreform; 65)*
- Wichert, P. von (2000): "Lust und Last der Lehre – Lehrevaluation in der Medizin". In: Forschung & Lehre, 7. Jg., 2000, H. 9, S. 474-476*

Wilkinson, J. (1992): "Der Harvard-Report – Eine Universität stellt sich auf den Prüfstand". In: Holtkamp, R.; Schnitzer, K. (1992): Evaluation des Lehrens und Lernens: Ansätze, Methoden, Instrumente – Evaluationspraxis in den USA, Großbritannien und den Niederlanden. Hannover: HIS GmbH 1992, S. 93-97

Wissenschaftsrat (1993): 10 Thesen zur Hochschulpolitik. Köln 1993

Anschrift des Verfassers:

Prof. Dr. Hans-Dieter Daniel
Wissenschaftliches Zentrum für Berufs- und Hochschulforschung
Universität Gesamthochschule Kassel
Henschelstr. 4
34109 Kassel