

Natalija el Hage

Studentische Urteile zur Lehr- und Studienqualität

**Analysen zu ihrer Validität anhand der Indikatoren und Skalen
im Studierendensurvey**

Hefte zur Bildungs- und Hochschulforschung (19)

Arbeitsgruppe Hochschulforschung, Sozialwissenschaftliche Fakultät,
Universität Konstanz, Dezember 1996

Natalija el Hage

Studentische Urteile zur Lehr- und Studienqualität

**Analysen zu ihrer Validität anhand der Indikatoren und Skalen
im Studierendensurvey**

Hefte zur Bildungs- und Hochschulforschung (19)

Arbeitsgruppe Hochschulforschung, Sozialwissenschaftliche Fakultät,
Universität Konstanz, Dezember 1996

Herausgeber der Reihe „Hefte zur Bildungs- und Hochschulforschung“:

Arbeitsgruppe Hochschulforschung, Universität Konstanz,
Fachbereich Geschichte und Soziologie, 78457 Konstanz
Tel. 07531/88-2896

Die AG Hochschulforschung im Internet:

<http://www.uni-konstanz.de/FuF/SozWiss/fg-soz/ag-hoc>

ISSN 1616-0398

Vorwort

In den Erhebungen des Studierendensurveys zur Studiensituation und zu den studentischen Orientierungen haben Fragen der Studien- und Lehrqualität seit dem Beginn 1983 eine Rolle gespielt. Selbstverständlich beziehen sich die Stellungnahmen der Studierenden im Rahmen des Studierendensurveys nicht auf einzelne Veranstaltungen und Lehrende, sondern auf die Situation und Verhältnisse im Fachbereich insgesamt.

Um zu klären, inwieweit die studentischen Antworten auf die verschiedenen Fragen und Vorgaben zutreffende Schlüsse auf die Studien- und Lehrqualität zulassen, sind zwei methodische Prüfungen notwendig. Zum einen ist zu prüfen, in welcher Weise die einzelnen Items zusammenhängen, eine „Dimension“ bilden, und welche „Meßqualität“ die Skalen erreichen. Zum anderen bedarf es der Prüfung, ob die Urteile der Studierenden von bestimmten Merkmalen und Gegebenheiten abhängen - z.B. dem Geschlecht, der Semesterzahl, dem Leistungsstand oder den eigenen Ansprüchen.

Über die Dimension und Skalen der insgesamt 52 Indikatoren (Items zur Lehr- und Studienqualität im Studierendensurvey) ist bereits von Frank Multrus ein ausführlicher Bericht vorgelegt worden. Er vermittelt zugleich eine Reihe wichtiger inhaltlicher Befunde über den Vergleich der Studienfächer und die Bedeutung der Kontakte und Kommunikation für eine gute Studienqualität insgesamt (vgl. Nr. 12 der Hefte zur Bildungs- und Hochschulforschung).

Der nun vorliegende Bericht von Natalija el Hage geht Fragen zur inhaltlichen Gültigkeit (Validität) der Indikatoren und Skalen nach. Sie knüpft dabei an oft geäußerte Vorbehalte an, die verschiedene Verzerrungen in den studentischen Urteilen vermuten. Diese Argumente behandelt sie Punkt für Punkt, wobei sie die Berechtigung dieser Vorbehalte anhand eines breiten Spektrums von Merkmalen und Faktoren untersucht. In ihre Schlußfolgerungen bezieht sie andere empirische Studien zum jeweiligen Problemkreis mit ein.

Zwei wichtige Resultate ihrer Aufarbeitungen seien hervorgehoben: Erstens treffen Unterstellungen offenbar nicht zu, die studentischen Aussagen seien wegen mangelnder Erfahrungen und Sachkenntnis oder wegen fehlender Reife kaum brauchbar. Zweitens zeigen sich durchaus manche Unterschiede in den Beurteilungen nach den Studienphasen, dem studentischen Fachinteresse und Leistungsstand, aber auch nach dem Grad der „Überfüllung“. Den Angaben der Studierenden kann demnach durchaus vertraut werden, mit der Prämisse, daß es sich um studentische Erfahrungen und Sichtweisen handelt. Messungen zur Studien- und Lehrqualität sind aber differenziert vorzunehmen, da Ansprüche, Interesse, Motivation und Studiengang eine Rolle spielen, welche Erfahrungen gemacht und welche Urteile abgegeben werden.

Die Analysen dieses Berichtes stützen sich zwar auf die Daten des Studierendensurveys, gehen aber in der Absicht darüber hinaus. Er ist zu verstehen als ein grundsätzlicher Beitrag zur Frage der Validität studentischer Veranstaltungskritik. Die Befunde erlauben die Folgerung, daß Evaluationen von Studium und Lehre gültig und zuverlässig über studentische Befragungen zu leisten sind, es aber wichtig ist, bei Vergleichen und Folgerungen bestimmte Faktoren zu beachten. Der Bericht soll dazu beitragen, manche Mißverständnisse und Vorbehalte auszuräumen, sowie die Brauchbarkeit und Verwendung studentischer Befragungen genauer zu bestimmen.

Tino Bargel

TABELLENVERZEICHNIS.....	IV
1 FRAGESTELLUNG UND DATENGRUNDLAGE.....	1
2 VERMUTETE EINFLÜSSE AUF STUDENTISCHE URTEILE ZUR LEHRE.....	4
3 INDIKATOREN UND SKALEN ZUR STUDIENQUALITÄT IM STUDIERENDENSURVEY.....	6
4 MANGELNDE SACHKENNTNIS DER STUDIERENDEN.....	9
<i>Schlußfolgerung</i>	15
5 UNREIFE UND UNERFAHRENHEIT DER STUDIERENDEN.....	16
<i>Schlußfolgerung</i>	17
6 DAS GESCHLECHT DER STUDIERENDEN BEEINFLUßT DIE BEWERTUNG.....	18
<i>Schlußfolgerung</i>	20
7 UNTERSCHIEDLICHES INTERESSE AM FACH, AN STUDIUM UND WISSENSCHAFT.....	21
<i>Schlußfolgerung</i>	26
8 UNTERSCHIEDLICHE LEISTUNGSBEREITSCHAFT.....	27
<i>Schlußfolgerung</i>	29
9 NOTEN UND LEISTUNGSFÄHIGKEIT.....	30
<i>Schlußfolgerung</i>	38
10 ÜBERFÜLLUNG VON LEHRVERANSTALTUNGEN.....	39
11 ZUSAMMENFASSUNG UND FOLGERUNGEN.....	40
LITERATUR	43
ANHANG	47

Beurteilung von Interaktion, Didaktik und Struktur im Studiengang und ...

Tabelle 1	
Inhaltlicher Zusammenhang einer vorhergehenden Ausbildung mit dem Studienfach.....	11
Tabelle 2	
Hochschulwechsel, Auslandsstudium und Mitarbeit in Forschungsprojekten	12
Tabelle 3	
Beschäftigung als studentische Hilfskraft oder Tutor.....	12
Tabelle 4	
Fachsemester der Studierenden an Universitäten	13
Tabelle 5	
Fachsemester der Studierenden an Fachhochschulen	13
Tabelle 6	
An Universitäten vor und nach Ablegen einer Zwischenprüfung	14
Tabelle 7	
Geschlecht der Studierenden.....	19
Tabelle 8	
Ausgewählte Fächergruppen nach Geschlecht an Universitäten	19
Tabelle 9	
Fächer der Ingenieurwissenschaften nach Geschlecht.....	20
Tabelle 10	
Wichtigkeit des speziellen Fachinteresses bei der Studienfachwahl	23
Tabelle 11	
Bedeutung des Studiums für die derzeitige Lebenssituation der Studierenden.....	24
Tabelle 12	
Wichtigkeit von Wissenschaft und Forschung	25
Tabelle 13	
Übereinstimmung mit dem „Wissenschaftlertypus“ nach GAWATZ.....	26
Tabelle 14	
Leistungsanspruch der Studierenden	28
Tabelle 15	
Studiertypen	29
Tabelle 16	
Übereinstimmung von persönlicher Leistungseinschätzung und erhaltenen Noten.....	32

Tabelle 17	
Korrelationstableau: Durchschnittsnote der Zwischenprüfung und Qualität von Interaktion, Didaktik und Struktur	33
Tabelle 18	
Abiturnote (Durchschnitt).....	34
Tabelle 19	
Persönliche Belastung durch Leistungsanforderungen im Fachstudium	35
Tabelle 20	
Sorgen bezüglich der Bewältigung des Studiums.....	36
Tabelle 21	
Subjektive Charakterisierung des Faches durch hohe Leistungsnormen und -ansprüche	37
Tabelle 22	
Korrelationstableau: Überfüllung der Hochschule und Qualität von Interaktion, Didaktik und Struktur	39

1 Fragestellung und Datengrundlage

Lehrqualität und Lehrevaluation haben sich in den letzten Jahren in der Hochschulpolitik immer stärker zu relevanten Themen entwickelt. In diesem Zusammenhang wird häufig die Frage nach geeigneten Instrumenten und Indikatoren zur Bestimmung von Qualität und Effizienz in der Lehre gestellt. Bei Vorhaben der Lehrevaluation wird meist zuerst an das Instrument der 'studentischen Veranstaltungskritik' gedacht. Die Beurteilung von Lehrenden durch Studierende ist jedoch immer noch umstritten.

In Band 13 der Hefte zur Bildungs- und Hochschulforschung („Zur Validität studentischer Veranstaltungskritiken“, el Hage, 1995) wurde daher auf der Grundlage einer Vielzahl empirischer Studien diskutiert, welche Zweifel gegenüber studentischen Veranstaltungskritiken begründet sind und welche nicht. Diese Aufarbeitung empirischer Ergebnisse konnte natürlich nicht alle Fragen endgültig beantworten. Entsprechend dem Stand der empirischen Forschung ist es aber möglich, den Schluß zu ziehen, daß studentische Urteile über Lehrveranstaltungen und Dozenten prinzipiell als gültige, zuverlässige und praktikable Grundlagen der Lehrevaluation herangezogen werden können. Dennoch bleiben an einigen Punkten offene Fragen.

In diesem Bericht sollen auf der Grundlage eines großen Datensatzes einige Bedenken gegenüber studentischer Veranstaltungskritik genauer untersucht und einige ungeklärte Probleme aufgegriffen werden. Grundlage dafür sind die Erhebungen des vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF, früher BMBW) geförderten Projektes „Entwicklung der Studiensituation und studentischer Orientierung“ der AG Hochschulforschung an der Universität Konstanz (zur Konzeption des Projekts vgl. Peisert, Bargel und Framheim, 1988).

Seit 1982 werden im Rahmen dieses Projektes in regelmäßigen Abständen (ca. alle 2-3 Jahre) repräsentative Daten von Studierenden erhoben, die Auskunft zur Verteilung und Veränderung ihrer Studiensituation, Erfahrungen und Orientierungen geben.

Das Erhebungsinstrument des Studierenden surveys beinhaltet ca. 120 meist geschlossene Fragen mit insgesamt ungefähr 660 Variablen, die sich sowohl auf die berufliche und gesellschaftliche Orientierung der Studierenden beziehen, als auch Auskunft zu Motivation und Studienverhalten geben. Darüber hinaus sind eine Reihe von Fragen enthalten, die die studentische Einschätzung der Lehrsituation und Studienqualität erfassen sollen.

Bisher wurden sechs Befragungen durchgeführt, die letzte im WS 94/95 (vgl. Bargel, Multrus & Ramm, 1996). Seit dem WS 92/93 werden auch Studierende ostdeutscher Hochschulen befragt. In die Stichprobe werden nach dem Zufallsprinzip ausgewählte Studierende aus ins-

gesamt 22 Hochschulen einbezogen (8 Universitäten und 6 Fachhochschulen in den alten Ländern und 5 Universitäten und 3 Fachhochschulen in den neuen Ländern). Im WS 94/95 wurden 19.644 Fragebogen versandt; die Rücklaufquote von 8.461 Fragebogen entspricht einer Beteiligung von 43 Prozent. Damit liegt ein weitgehend repräsentativer Querschnitt der deutschen Studierendenschaft vor.

Seit Beginn des Projekts sind Fragen über Studienqualität und Lehrevaluation zentraler Bestandteil der Untersuchungen. Die Auswahl der Fragen und Vorgaben zu diesem Themenbereich stützte sich auf damals vorliegende Untersuchungen zum Hochschulunterricht (Keil/Piontkowski 1973), zur Organisation der Lehre (Ipsen/Portele 1976; Oehler u.a. 1978), sowie zu verschiedenen Ausarbeitungen aus dem Umkreis der Hochschuldidaktik und der Lehr-Lernforschung. Außerdem wurden eigene frühere Untersuchungen zur Hochschulsozialisation und zur Studiensituation herangezogen (Framhein u.a. 1980; Gleich u.a. 1982).

Mit der einsetzenden Evaluation der Lehre an deutschen Hochschulen, zumeist orientiert an amerikanischen Vorbildern, konnten aus diesem Bereich Ansätze und Erhebungen aufgenommen werden (Preißer 1991; Holtkamp/ Schnitzer 1992; Webler u.a. 1993). Schließlich wurde auch darauf geachtet, mit den Fragen und Vorgaben des Studierendensurveys den Vergleich zu evaluativen Rankings herstellen zu können, ohne daß freilich selber solche Hochschul-Rankings beabsichtigt waren (vgl. SPIEGEL 1990 und 1993; RCDS 1992).

Die Fragen im Studierendensurvey beziehen sich nicht auf einzelne Veranstaltungen, sondern verlangen die Einschätzung der insgesamt erfahrenen Studiensituation und empfundenen Lehrqualität im Studienfach. Die studentischen Urteile sind daher Aussagen zur Beurteilung der Lehre in dem jeweiligen Fachbereich und nicht, wie sonst meist der Fall, die Bewertung einzelner Lehrender oder einzelner Veranstaltungen.

Für die Fächer Medizin und Rechtswissenschaften wurden jeweils Fachmonographien erstellt (Bargel & Ramm, 1994 und Bargel, Multrus & Ramm, 1996). Sie stellen für die beiden Fächer eine hochschulübergreifende Berichterstattung bezüglich der Studien- und Lehrqualität aus Sicht der Studierenden dar. Gegenwärtig wird eine Fachmonographie der Geisteswissenschaften erstellt, wobei die Fächer Germanistik, Anglistik und Geschichte einbezogen werden. Im Sinne eines Lehrberichts wird das Augenmerk vor allem auf strukturelle, allgemeine Probleme des jeweiligen Faches gelenkt, wie es die Studierenden erleben und beurteilen; weniger interessieren mögliche Mängel einzelner Hochschulen oder die Aufstellung eines Hochschul-Rankings.

Die Beurteilung einzelner Kurse und Lehrender wird meist unter dem Schlagwort 'studentische Veranstaltungskritik' zusammengefaßt. Von seiten der Lehrenden werden häufig Bedenken gegen die Zuverlässigkeit und Gültigkeit dieses Instrumentes der Lehrevaluation geäußert

(z.B. Süllwold, 1992). Dabei wird angenommen, daß bestimmte Merkmale der Studierenden, wie z.B. Geschlecht oder Alter, ihre Urteile beeinflussen und daher kaum zuverlässige Aussagen bezüglich der tatsächlichen Lehrqualität zulassen.

Inwieweit können solche Bedenken durch Ergebnisse aus einer Studie beurteilt werden, die auf die Bewertung des erfahrenen Gesamtlehrangebots und der Studienqualität im entsprechenden Semester abzielt und nicht die Lehrbewertungen einzelner Kurse ermittelt?

Geht man davon aus, daß am ehesten globale, summative Antworten mit einem Bias behaftet sind, d.h. aufgrund unerwünschter Einflußfaktoren verzerrt werden, wie es Marsh (1982) in bezug auf studentische Veranstaltungskritik nachweisen konnte, so kann folgendes angenommen werden: Mögliche verzerrende Einflußfaktoren auf die wahrgenommene Lehrqualität müßten bei kursübergreifenden Fragebogen, der sich nicht einmal auf einzelne Lehrveranstaltungen, sondern auf ihre Gesamtheit bezieht, noch weitaus stärker wirksam und damit wie durch ein Vergrößerungsglas wahrgenommen werden können.

Für die alternative Hypothese, daß unerwünschte Einflußfaktoren bei solchen Antworten nicht zum Tragen kommen, jedoch in Fragebogen, die auf einzelne Kurse bezogen sind, unerwünschte Verzerrungen hervorrufen, gibt es weder empirische noch theoretische Anhaltspunkte.

Daher wird davon ausgegangen, daß störende Einflüsse, die im Studierendenurvey im Zusammenhang mit der Beurteilung der Studien- und Lehrqualität gefunden werden, sich ebenfalls in relativ grob gestrickten Fragebogen zu studentischen Veranstaltungskritiken wiederfinden lassen. Je detaillierter ein Fragebogen ist, desto weniger wahrscheinlich sind solche unerwünschten Verzerrungen (vgl. Marsh, 1982).

Weiter ist davon auszugehen, daß Faktoren, die sich bei den Fragen des Studierendenurveys nicht als Bias herausstellen, d.h. bei denen kein Zusammenhang mit den Beurteilungen zur Lehrqualität zu finden ist, sich auch in einfachen Fragebogen der studentischen Veranstaltungskritik nicht verzerrend auswirken.

Als Grundlage der Überprüfung dieser Fragestellung werden Ergebnisse, des von der AG Hochschulforschung verwendeten Erhebungsinstrumentes (des Studierendenurveys), verwendet. Dabei wurden die im Fragebogen enthaltenen Angaben zur Lehrqualität, zur Betreuung und Beratung, sowie zu Inhalt, Didaktik und Struktur des Studiengangs benutzt.

Multrus (1995) hat die entsprechenden Items - insgesamt 52 - einer Faktorenanalyse unterzogen. Dieses Verfahren ermöglicht, ähnlich gelagerte Antworten zu identifizieren und dadurch zugrundeliegende Dimensionen sichtbar zu machen. Diese 'Dimensionen der Lehrqualität' stellen eine Grundlage dieses Berichts dar. Anhand dieser Dimensionen wird im folgenden

überprüft, inwieweit sich unerwünschte Einflüsse (Geschlecht, Alter etc.) auf die Lehrbeurteilung auswirken.

Die nachfolgenden Auswertungen und Analysen haben ein doppeltes Ziel: Zum einen liefern sie unmittelbare Einsichten in Zusammenhänge zwischen Merkmalen auf studentischer Seite und deren Urteile zur Lehrsituation oder Studienqualität; dadurch ermöglichen sie Schlüsse auf die Validität der benutzten Fragen und Items. Zum anderen dienen sie dazu, einige Datenlücken zu schließen, wobei vor allem jene vermuteten Einflüsse auf studentische Stellungnahmen behandelt werden, bei denen die Datenlage schwach ist oder aber Besonderheiten des deutschen Hochschulsystems vorliegen, zu denen die meist nordamerikanischen Untersuchungen keine oder kaum Aussagen treffen.

2 Vermutete Einflüsse auf studentische Urteile zur Lehre

Obwohl seit einigen Jahren deutsche Hochschullehrer und -lehrerinnen Erfahrungen mit studentischer Veranstaltungskritik gewonnen haben, hat ihre Skepsis gegenüber solchen Verfahren kaum abgenommen. Auf Tagungen und Kongressen, in Senaten und Kommissionen, gelegentlich auch in Zeitungsartikeln werden noch immer Bedenken geäußert.

Als Beispiel für häufig genannte Zweifel seien die von Süllwold (1992) vermuteten Einflußgrößen genannt, welche die Lehrbewertung verzerren könnten: Fachsemester, geschlechtsspezifische Zusammensetzung, voruniversitäre Bildung, Intelligenzniveau, Fleiß, Leistungsmotivation und ideologische Bedürfnisse.

Hinzu kommen weitere Einwände, die ebenfalls in Veröffentlichungen aus anderen Ländern zu finden sind. So nennt Aleamoni (1987) als häufigste Einwände gegen studentische Veranstaltungskritik:

- I. Studierende können die Lehre nicht konsistent beurteilen, da sie folgende Eigenschaften aufweisen: 'Unreife, fehlende Erfahrung und Launenhaftigkeit'.
- II. Nur Kollegen/ Kolleginnen können die Lehre entsprechend kompetent bewerten.
- III. Die meisten Bewertungsbögen sind nichts weiter als ein 'Beliebtheitswettbewerb'.
- IV. Studierende können die Lehre nicht bewerten, bis sie nicht genügend Distanz zum Kurs bzw. zur Lehranstalt besitzen.
- V. Bewertungsbögen sind unreliabel und nicht valide.
- VI. Externe Variablen (z.B. Kursgröße) beeinflussen die Bewertung.
- VI. Noten (erwartete oder erhaltene) beeinflussen die Bewertung.

Damit ist ein großer Teil der Argumente, die immer wieder gegen die Beurteilung der Lehrqualität durch Studierende hervorgebracht werden, benannt. Eine Diskussion dieser Argumente muß sich bislang vor allem auf Studien stützen, die zum großen Teil aus den Vereinigten Staaten stammen (vgl. el Hage, 1996).

Anhand dieser Studien läßt sich zwar klären, daß Fragebogen der studentischen Veranstaltungskritik so konstruiert werden können, um den üblichen methodischen Kriterien an Zuverlässigkeit und Gültigkeit zu genügen. Aber es war nicht möglich, den vermuteten verzerrenden Einfluß einiger Variablen endgültig zu klären, da manche der empirischen Untersuchungen selbst gewisse Mängel aufweisen und zu bestimmten Themen nur sehr wenige Studien vorliegen.

Außerdem konnten, bis auf wenige Ausnahmen, keine deutschen Studie einbezogen werden. Die Situation an den deutschen Hochschulen unterscheidet sich in mancher Hinsicht jedoch wesentlich von der an nordamerikanischen Universitäten. Daher konnten auch zu bestimmten typisch deutschen Besonderheiten keine bzw. wenig Aussagen getroffen werden. Dies gilt z.B. für die Frage, ob sich die „Überfüllung“ vieler Fächer auf die Beurteilung der Lehre auswirkt, oder für die Frage, ob Studierende mit sehr unterschiedlichen Vorerfahrungen (z.B. mit abgeschlossener beruflicher Ausbildung) die Lehrqualität insgesamt anders einschätzen.

Einige dieser noch offenen oder nur z.T. geklärten Fragen zur Gültigkeit und Genauigkeit studentischer Beurteilungen werden in dem vorliegenden Bericht behandelt. Aufgegriffen werden vor allem grundsätzliche Einwände gegen die Gültigkeit und damit Verwendbarkeit studentischer Veranstaltungskritik, die in erster Linie die mangelnde Sachkenntnis, Unerfahrenheit und Unreife der Studierenden betonen.

Häufig wird in Diskussionen auch vorgebracht, daß Merkmale der Studierenden dazu führen, die Lehre unterschiedlich wahrzunehmen wird. Dabei werden als Merkmale meist Geschlecht, Interesse am Studium und Leistungsmotivation genannt.

Weiter wird der Vermutung nachgegangen, ob Leistungsbeurteilungen, die die Studierenden erhalten, sich maßgeblich auf die Beurteilungen der von ihnen erfahrenen Lehrqualität auswirken. Schließlich wird versucht zu ermitteln, inwieweit sich die Überfüllung der Hochschulen auf die Beurteilung der Lehrqualität auswirkt.

Die aufgegriffenen Argumente umfassen zentralen Kritikpunkte gegen die Verwendung studentischer Veranstaltungskritiken, die sich hauptsächlich auf unerwünschte Einflüsse aufgrund externer Einflußfaktoren beziehen. Methodische Probleme anderer Art, z.B die Vergleichbarkeit von Bewertungen aus verschiedenen Kursen, die Heterogenität studentischer Urteile und ähnliches, werden in diesem Bericht nicht weiter thematisiert (vgl. dazu el Hage, 1996).

Übersicht 1

Argumente gegen die Validität studentischer Veranstaltungskritiken

1. Mangelnde Sachkenntnis der Studierenden

Studierenden fehle die Sachkenntnis, um die Kompetenz der Lehrenden und die Lehrqualität überhaupt angemessen einschätzen zu können.

2. Unerfahrenheit und Unreife der Studierenden

Studierenden fehle es an Erfahrung und zum Teil am notwendigen Ernst. Diese Unreife führe zu beliebigen und beeinflussbaren Urteilen.

3. Das Geschlecht beeinflusst die Bewertung

Studentinnen würden die Lehre anders 'empfinden' als ihre Kommilitonen. Daher würden sie auch Lehraspekte anders beurteilen.

4. Unterschiedliches Interessen am Fach, an Studium und Wissenschaft

Studierende mit unterschiedlich ausgeprägten Interessen kämen zu anderen Beurteilungen; daher hänge die 'Lehrqualität' weniger vom Lehrenden als von den Studierenden ab.

5. Unterschiedliche Leistungsbereitschaft

Unterschiedlich leistungsbereite Studierende, unterschiedlich leistungsmotiviert, würden Lehre unterschiedlich wahrnehmen und bewerten.

6. Noten und Leistungsfähigkeit

Die Leistungsfähigkeit der Studierenden und die Strenge der Benotung würden deren Lehrbeurteilungen maßgeblich bestimmen.

7. Überfüllung beeinträchtigt die Bewertung

Volle Veranstaltungen und überfüllte Fachbereiche würden insgesamt schlechter beurteilt.

3 Indikatoren und Skalen zur Studienqualität im Studierendensurvey

Das verwendete Instrument, der Studierendensurvey, beinhaltet eine Reihe von Fragen zur Lehrqualität, die aufgrund theoretischer Vorüberlegungen zusammengestellt wurden. In diesem Fragebogen sind zehn verschiedene Komponenten der Studienqualität enthalten, wobei allerdings ihre Erfassung in unterschiedlicher Breite erfolgt ist (vgl. Bargel, 1995):

- fachliche Qualität des Studien- und Lehrangebotes,
- strukturelle Qualität des Studienganges,
- didaktische Qualität in den Lehrveranstaltungen,
- die interaktiv-tutoriale Qualität der Betreuung und Beratung,
- Höhe und Umfang der fachlichen Leistungsanforderungen,
- überfachliche Anforderungen an Kommunikation und Kompetenzen,
- Ausmaß an Praxisbezug,
- Ausmaß an Forschungsbezug,
- Studierertrag in der fachlich-beruflichen Qualifizierung,
- Förderung allgemeiner Kompetenzen im Sinne von 'Schlüsselqualifikationen'.

Multrus (1995) sonderte 52 Einzelitems des Studierendensurveys aus, die auf die Erfassung der Studien- und Lehrqualität in den einzelnen Fachbereichen abzielen. Mit Hilfe einer Faktorenanalyse konnte er bei diesen Items insgesamt sechs einzelne Dimensionen identifizieren: (1) Lehrqualität und Interaktion, (2) Arbeitsanforderung im Studium, (3) Förderung von allgemeinen Fähigkeiten, (4) allgemeinen Anforderungen im Studium, (5) strukturelle Qualität, (6) Praxis- und Berufsbezug. Die erste Skala konnte nach Eliminierung trennschwacher Items in zwei Subskalen unterteilt werden, zum einen zur didaktisch-inhaltlichen Qualität und zum anderen zur interaktiv-tutorialen Qualität.

Für die vorliegende Untersuchung wurde nur mit jenen drei Skalen gearbeitet, die folgende zentrale Bereiche der Lehrqualität erfassen:

- Betreuung und Beratung,
- Didaktik und Inhalt,
- sowie Struktur und Aufbau des Studienganges.

Daß gerade diese Dimensionen ausgewählt wurden, liegt an der Fragestellung: Inwieweit wirken sich unerwünschte Einflüsse auf die Beurteilung der Lehrqualität aus? Daher wurden nur jene Dimensionen in die Auswertung miteinbezogen, die in ähnlicher Form zumeist in Fragebogen zur studentischen Lehrveranstaltungskritik zu finden sind (vgl. z. B. Diehl, 1994 oder Rindermann & Amelang, 1994).

Es handelt sich um die Skalen: zur didaktisch-inhaltlichen Qualität (Variante A), zur interaktiv-tutorialen Qualität (Variante A) und zur strukturellen Qualität (Variante B). (Für die Entwicklung der Skalenmaße vgl. Multrus, 1995).

Die *Skala zur didaktisch-inhaltlichen Qualität* umfaßt sechs Fragen, die sich auf den Vortrag der Lehrenden und die Durchführung der Lehrveranstaltung beziehen, und zwar in konkreterer und allgemeinerer Art:

Charakteristika: Lehrende sind auf Veranstaltungen gut vorbereitet.

Lehrveranstaltung: Vortrag ist verständlich.

Lehrveranstaltung: Dozent vergewissert sich, daß Stoff verstanden ist.

Lehrveranstaltung: Beispiele der Dozenten fördern Lehrerfolge.

Lehrveranstaltung: Dozent bringt übersichtliche Zusammenfassungen.

Bewertungsaspekt: Art und Weise der Durchführung der Lehrveranstaltungen.

Die *Skala zur interaktiv-tutorialen Qualität* berücksichtigt fünf Aspekte der Betreuung und Beratung sowie des sozialen Klimas der Beziehungen:

Charakteristika: Gute Beziehung zwischen Studierenden und Lehrenden.

Situation in Fachbereich: Persönliche Beratung durch Lehrende, wenn nötig.

Situation in Fachbereich: Ergebnisse von Tests werden erläutert.

Situation in Fachbereich: Berücksichtigung von studentischen Anregungen.

Bewertungsaspekt: Betreuung und Beratung durch Lehrende.

Die *Skala zur strukturellen Qualität* schließlich beinhaltet fünf Fragen zur Gliederung und Aufbau eines Studienganges.

Charakteristika: Gut gegliederter Studienaufbau.

Bewertungsaspekt: Struktur des Studienganges.

Charakteristika: Klarheit der Prüfungsanforderungen.

Situation im Fachbereich: Angekündigter Stoff wird innerhalb des Semester vermittelt.

Situation im Fachbereich: Lehrveranstaltungen sind ausgefallen.

Diese drei Dimensionen wurden für die Untersuchung herangezogen. Im folgenden wird die Skala zur interaktiv-tutorialen Qualität kurz 'Interaktion' genannt, die Skala zur didaktisch-inhaltlichen Qualität mit 'Didaktik' abgekürzt und die Skala zur strukturellen Qualität als 'Struktur' bezeichnet.

Um zu klären, ob verschiedene Faktoren die Beurteilung der Lehrqualität beeinflussen oder verzerren, wird folgendermaßen vorgegangen: Die aufgeführten Argumente wurden Items des Studierenden-surveys gegenübergestellt. Bei einfachen Merkmalen, wie z.B. dem Geschlecht, konnte direkt verglichen werden, ob Frauen die Lehre anders beurteilen als Männer. Anders dagegen mußte bei abstrahierten Thesen vorgegangen werden, wie z.B. daß Studierende unreif und unerfahren seien und daher die Lehre nicht beurteilen könnten. Bei solchen Argumenten mußte erst überprüft werden, wie sich Begriffe wie „Unreife“ oder „Unerfahrenheit“ überhaupt operationalisieren lassen, d.h. welche der verfügbaren Angaben darüber Auskunft geben könnten. Die entsprechenden Merkmale der Studierenden wurden dann in Beziehung zu den drei ausgewählten Dimensionen der Studien- und Lehrqualität gesetzt.

Da sich aufgrund der hohen Zahl befragter Studierender in der Stichprobe des Studierenden-surveys die meisten Vergleiche mit Hilfe der klassischen Signifikanztests selbst bei minimalen Differenzen signifikant voneinander unterschieden hätten, wurde auf solche Tests verzichtet. Verwendet wird stattdessen eine von Lind (1993) vorgeschlagene Bestimmung, bei der als Maß der Effektstärke bei Survey-Studien die relative Differenz zwischen den Meßwerten verwendet werden sollte.

Die Übereinkunft darüber, wie groß diese relative Differenz sein sollte, um von wirklichen Unterschieden sprechen zu können, ist zwar methodisch noch nicht getroffen worden. Allerdings existieren gewisse psychometrische Gepflogenheiten. Mit Lind schließen wir uns der 'faktischen Standardisierung auf konventioneller Basis' an und sehen alle Differenzen, die größer als 5 Prozentpunkte sind als Hinweis auf eine unterschiedliche Bewertung an. Diese werden dann, wie im statistischen Sprachgebrauch häufig, als 'bedeutsame' oder 'tendenzielle' Unterschiede bezeichnet. Bei Differenzen von mehr als 10 Prozent wird davon ausge-

gangen, daß eindeutig unterschiedliche Bewertungen vorliegen und sie daher 'sehr bedeutsam' oder 'signifikant' sind.

Aufgrund der gewählten Vorgehensweise ergibt sich für den vorliegenden Bericht folgende Gliederung: Die Argumente gegen studentische Veranstaltungskritik werden einzeln der Reihe nach erörtert. Es werden, falls vorhanden, jeweils erst die Befunde anderer Studien aufgeführt und dann die Ergebnisse auf der Grundlage des Studierendensurveys berichtet. Abschließend erfolgt dann eine zusammenfassende Einschätzung, inwieweit sich die untersuchten Faktoren tatsächlich auf die Beurteilungen der Lehrqualität auswirken und somit die Annahme einer 'Verzerrung' berechtigt oder unberechtigt ist.

4 Mangelnde Sachkenntnis der Studierenden

Eines der meistgenannten Argumente gegen studentische Veranstaltungskritiken unterstellt, daß Studierende gar nicht die nötige Sachkenntnis besitzen würden, um die Qualität der Lehre tatsächlich beurteilen zu können. Dabei steht im Vordergrund, daß eine Beurteilung der Lehrqualität vor allem eine Beurteilung des Lehrinhalts ist. In Fragebogen der studentischen Veranstaltungskritik spielt eine Beurteilung des Inhalts jedoch meist kaum eine Rolle. In seltenen Fällen ist gelegentlich die eine oder andere Frage dazu zu finden.

Meist wird vor allem die Art und Weise der Stoffpräsentation, die Nützlichkeit der Veranstaltung für Prüfungen und ähnliches bewertet. Natürlich kann dagegen eingewandt werden, daß Studierende auch in der Beurteilung von eher didaktischen Aspekten der Lehre keine ausgebildeten Fachleute sind. Dennoch ist es berechtigt zu diskutieren, ob sie in irgendeiner Weise nicht doch Experten und Expertinnen in diesem Bereich sind. Auch gibt die Studie von McKeachie und Lin (1978) einige Hinweise darauf, daß die Lehrbeurteilungen von Studierenden und trainierten Beobachtern gut übereinstimmen.

Weiter ist es möglich zu vergleichen, ob Studierende tatsächlich die Lehre anders einschätzen als Personen, die als Experten und Expertinnen in dem jeweils zu beurteilenden Bereich gesehen werden: nämlich die Lehrenden selbst. Eine der wohl umfangreichsten Untersuchungen zu dieser Fragestellung ist von Marsh (1982) durchgeführt worden. Er überprüfte die Übereinstimmung zwischen den Selbstbeurteilungen von Dozenten und Dozentinnen und der Kursbewertung durch Studierende. Die Ergebnisse dieser Korrelationsuntersuchung ergaben kongruente Werte. Dies ist als Hinweis darauf zu verstehen, daß Studierende, die über wenig(er) Sachkenntnis verfügen, die Lehrqualität dennoch sehr ähnlich einschätzen, wie Lehrende, von denen anzunehmen ist, daß sie über viel Sachkenntnis verfügen.

Immer wieder wird zudem vermutet, daß niedrigere Semester ihre Lehrenden signifikant schlechter bewerten. Danielsen und White (1976), Patricia B. Elmore und Pohlmann (1978), Hornbostel und Daniel (1993), Daniel (1995) ebenso wie Rindermann (1996) konnten jedoch keine signifikanten Unterschiede in der Lehrbewertung unterschiedlicher Semester finden. Weder Blount et al. (1978) noch Felicita F. Romeo und Weber (1985) fanden Unterschiede in der Veranstaltungskritik von 'graduate' und 'ungraduate' Studierenden. In der Studie von Bruton und Crull (1982) hatte die Anzahl der Studienjahre einen geringen, aber doch signifikant negativen Einfluß auf die Kursbewertung. Adriane Gaffuri et al. (1982) konnten nur eine Durchschnittskorrelation von .18 zwischen einem Item und dem Kurslevel finden.

Wir wollen anhand der Daten des Studierendensurveys überprüfen, ob Studierende mit geringerer Sachkenntnis Lehre anders einschätzen, als Studierende, von denen anzunehmen ist, daß sie mehr inhaltliche Kenntnisse und/oder Lehrerfahrung haben. Dafür ist es vorab notwendig, den Faktor 'Sachkenntnis' zu operationalisieren.

Welche Faktoren können dazu führen, daß Studierende zum Inhalt des Faches, zur Stoffpräsentation, zur Betreuung und Beratung durch die Lehrenden und zum Studienaufbau mehr Sachkenntnisse besitzen? Diese können bei Studierenden eigentlich nur aufgrund einer höheren eigenen Erfahrung erworben worden sein. Im Studierendensurvey sind einige Fragen enthalten, die auf eine höhere Erfahrung hinweisen, d.h. eine höhere „Sachkenntnis“ erwarten lassen (vgl. Übersicht 2).

Übersicht 2

Studierende können höhere Sachkenntnis erworben haben durch:

1. Eine vor-universitäre Ausbildung, die fachlich mit dem Studium zusammenhängt.
2. Hochschulwechsel.
3. Auslandsstudium.
4. Mitarbeit an einem Forschungsprojekt.
5. Tutoratsstelle.
6. Stelle als Hilfswissenschaftler, -wissenschaftlerin.
7. Verweildauer an der Hochschule.
8. Abschluß des Grundstudiums.

Die in Übersicht 2 angeführten acht Aspekte werden im weiteren Punkt für Punkt behandelt.

(1) Von Studierenden, die vor Beginn ihres Studiums eine *Ausbildung abgeschlossen* haben, die inhaltlich eng mit ihrem Studienfach zusammenhängt, ist anzunehmen, daß sie auch über mehr Sachkenntnisse in diesem Bereich verfügen. Daher könnte vermutet werden, daß diese

Studierenden die Lehrveranstaltungen an den Hochschulen unterschiedlich erleben und beurteilen könnten.

Dies ist jedoch offensichtlich nicht der Fall (vgl. Tabelle 1). Studierende, die bereits eine Ausbildung abgeschlossen haben, unterscheiden sich in der Beurteilung der Lehre nicht von ihren Kommilitonen und Kommilitoninnen, die gleich ein Studium begonnen haben. Dies deckt sich mit den Ergebnissen von Daniel (1995), der ebenfalls keinen bedeutsamen Einfluß auf die Lehrbeurteilungen ermitteln konnte.

Tabelle 1 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach dem inhaltlichen Zusammenhang einer vorhergehenden Ausbildung mit dem Studienfach (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)				
Dimensionen der Studienqualität	n = Untergruppen %	Zusammenhang Ausbildung - Studium		
		gar nicht / keine Ausbildung gemacht	nur wenig / teilweise	überwiegend / sehr eng
		5466 66,2	1143 13,9	1643 19,9
Interaktion	Mittelwert	47,6	47,1	48,1
	Standardabweichung	19,8	19,5	20,3
Didaktik	Mittelwert	51,4	50,1	52,4
	Standardabweichung	14,9	15,2	14,9
Struktur	Mittelwert	47,4	46,5	49,2
	Standardabweichung	14,5	14,7	14,0

Quelle: Konstanzer Projekt Studiensituation: Frage 10: „Hängt Ihre berufliche Ausbildung bzw. Ihre Berufstätigkeit mit Ihrem Studienfach inhaltlich zusammen?“

(2) Von Studierenden, die die *Hochschule gewechselt* oder sogar (3) im *Ausland studiert* haben, könnte ebenfalls vermutet werden, daß sie neue Erfahrungen sowohl in bezug auf die Strukturierung eines Studiengang als auch in anderen Bereichen gesammelt haben und deswegen möglicherweise anders, vielleicht kritischer urteilen.

Unsere Ergebnisse zeigen, daß weder ein Hochschulwechsel noch ein Auslandsstudium zu anderen Urteilen der Studierenden führt. Auch an der eigenen Hochschule ist es möglich mit einem Studienfach enger vertraut zu werden, in dem man z.B. (4) an einem *Forschungsprojekt mitarbeitet* oder eine Tutoren- oder Hilfskraftstelle annimmt.

Ebenfalls läßt die Mitarbeit in einem Forschungsprojekt, zumeist werden dazu erfahrenere Studierende herangezogen, keine anderen Urteile zur Interaktion, Didaktik und Studiengangsgliederung erkennen (siehe Tabelle 2).

Diese Befunde decken sich mit den Ergebnissen von Hornbostel und Daniel (1993), die diese in der zweiten SPIEGEL-Erhebung (vgl. SPIEGEL-SPEZIAL, 1993) sammeln konnten.

Tabelle 2 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach Hochschulwechsel, Auslandsstudium und Mitarbeit an Forschungsprojekten (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)							
Dimensionen der Studienqualität		Haben Sie im Laufe Ihres Studiums ...					
		die Hochschule gewechselt?		im Ausland studiert?		an Forschungsprojekten mitgearbeitet?	
		nein	ja	nein	ja	nein	ja
Unterguppen %		6972 83,6	1372 16,4	7844 93,8	522 6,2	7363 88,1	990 11,9
Interaktion	Mittelwert	47,6	47,6	47,6	46,6	47,5	48,4
	Standardabweichung	19,9	19,8	19,9	19,5	19,7	21,2
Didaktik	Mittelwert	51,6	50,3	51,4	51,0	51,4	51,1
	Standardabweichung	14,9	15,2	14,9	15,1	14,8	15,7
Struktur	Mittelwert	47,9	45,8	47,7	45,9	47,5	47,8
	Standardabweichung	14,4	14,7	14,4	14,7	14,4	14,8

Quelle: Konstanzer Projekt Studiensituation: Frage 47 / 48. (Frage wie Tabellentext)

Schließlich kommen jene Studierenden, die (5) eine *Tutorenstelle* oder (6) *Hilfskraftstelle* hatten, zu keiner anderen Einschätzung der Lehrqualität als die restlichen Studierenden (siehe Tabelle 3). Auch Hornbostel und Daniel (1993) konnten keine Unterschiede zwischen studentischen Hilfskräften und anderen Studierenden finden.

Tabelle 3 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach Beschäftigung als studentische Hilfskraft oder Tutor (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)					
Dimensionen der Studienqualität		Beschäftigung als			
		studentische Hilfskraft		Tutor	
n =		nein	ja	nein	ja
Unterguppen %		6820 83,2	1378 16,8	6978 92,1	596 7,9
Interaktion	Mittelwert	47,0	50,6	47,4	47,0
	Standardabweichung	19,8	20,2	19,9	18,7
Didaktik	Mittelwert	51,5	51,1	51,4	50,2
	Standardabweichung	14,9	15,0	14,9	14,8
Struktur	Mittelwert	47,6	47,0	47,6	48,1
	Standardabweichung	14,4	14,8	14,3	14,6

Quelle: Konstanzer Projekt Studiensituation: Frage 28: „Sind Sie oder waren Sie schon einmal als studentische Hilfskraft / Tutor beschäftigt?“

Der einfachste Indikator dafür, daß Studierende zu mehr Sachkenntnis gelangt sind, ist schlicht die *Dauer ihres Studiums* (7). Wenn der Vorwurf zutrifft, daß Studierende zu wenig Sachkenntnis haben, um die Güte von Lehrveranstaltungen einzuschätzen, müßten sie mit längerer Verweildauer an der Hochschule zu anderen Resultaten kommen. Die Befunde ande-

rer Studien, z.B. von Hornbostel und Daniel (1993), konnten diese Vermutung nicht bestätigen.

Tabelle 4 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach Fachsemestern: Universitäten (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)											
Dimensionen der Studienqualität	Fachsemester										
	1	2	3,4	5,6	7,8	9,10	11,12	13,14	15,16	mehr	
n =	740	176	1043	1050	1148	977	581	353	162	221	
Untergruppen %	11,5	2,7	16,2	16,3	17,8	15,1	9,0	5,5	2,5	3,4	
Interaktion	Mittelwert	50,6	48,0	47,9	46,2	46,3	45,2	43,7	41,5	44,6	42,7
	Standardabweichung	18,2	19,1	20,1	20,2	20,1	20,3	20,4	20,1	20,2	20,9
Didaktik	Mittelwert	54,8	50,6	52,1	51,4	50,0	50,4	48,8	47,1	48,8	46,2
	Standardabweichung	14,2	15,4	14,4	14,1	15,0	15,3	15,1	14,3	15,9	16,0
Struktur	Mittelwert	49,6	45,0	48,6	47,4	45,9	46,6	44,7	45,0	44,4	42,1
	Standardabweichung	14,0	13,0	13,6	14,3	14,9	15,1	15,0	13,9	15,9	15,3

Quelle: Konstanzer Projekt Studiensituation: Frage 52: „Im wievielten Fachsemester befinden Sie sich?“

Die Annahme, Studierende mit längerer Studiendauer würden die Lehre anders betrachten und bewerten, kann im Großen und Ganzen anhand der Daten des Studierendenurveys aber für die Universitäten und Fachhochschulen bestätigt werden. Die Beurteilung der Lehrqualität sinkt mit der Verweildauer an einer Hochschule gleich welcher Art (vgl. Tabelle 4 und Tabelle 5).

Möglicherweise spielen hierbei auch Selektionseffekte eine gewisse Rolle, die dazu führen, daß in höheren Semestern die Dimensionen der Studienqualität deutlich schlechter beurteilt werden. Eher ist aber anzunehmen, daß die Unzufriedenheit der Studierenden im Studienverlauf kumuliert, und daher mit den Fachsemestern die schlechtere Beurteilung steigt.

Tabelle 5 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach Fachsemestern: Fachhochschulen (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)											
Dimensionen der Studienqualität	Fachsemester										
	1	2	3,4	5,6	7,8	9,10	11,12	13,14	15,16	mehr	
n =	215	109	395	520	368	136	49	19	12	20	
Untergruppen %	11,7	5,9	21,4	28,2	20,0	7,4	2,7	1,0	0,7	1,1	
Interaktion	Mittelwert	58,9	53,9	54,7	50,4	50,9	49,8	49,9	46,5	35,7	42,3
	Standardabweichung	17,2	18,1	18,0	18,2	18,6	19,9	19,0	17,4	17,7	18,6
Didaktik	Mittelwert	58,9	56,7	55,6	52,4	51,2	51,2	48,5	47,8	43,2	48,7
	Standardabweichung	13,7	15,4	13,7	14,6	15,5	14,9	15,0	14,5	20,3	12,4
Struktur	Mittelwert	53,8	50,4	51,7	49,9	48,5	48,6	47,9	47,6	44,9	46,5
	Standardabweichung	13,8	12,9	12,6	13,7	13,8	12,7	14,8	15,7	19,6	17,3

Quelle: Konstanzer Projekt Studiensituation: Frage 52: „Im wievielten Fachsemester befinden Sie sich?“

Vergleicht man die Studienanfänger und -anfängerinnen mit den Studierenden des fünften bzw. sechsten Studienjahres, so ist insgesamt sogar eine statistisch bedeutsame Verschlechterung festzustellen. An den Fachhochschulen findet sich ein ähnliches Muster.

Am deutlichsten ist die Verschlechterung in der Lehrbewertung zwischen dem ersten und dritten bzw. vierten Semester zu bemerken. Bereits zwischen dem ersten Semester und dem vierten Studienjahr gibt es statistisch bedeutsame Unterschiede. Insgesamt kann also gefolgert werden, daß Studierende mit einer längeren Studiendauer die Lehre kritischer betrachten als Studienanfänger. Dieses Ergebnis erweist sich als diskrepant zu den aufgeführten Studien.

Dies kann möglicherweise auf die in Deutschland übliche Gliederung des Studiums in Grund- und Hauptstudium zurückgeführt werden. (8) Studierende, die ihr *Grundstudium abgeschlossen* haben, haben durchschnittlich gut ein Drittel ihres Studiums und die ersten umfangreichen Prüfungserfahrungen hinter sich. Sie verfügen somit über einen ganz anderen Erfahrungsschatz als z.B. Studienanfänger und -anfängerinnen.

Dennoch sind keine prinzipiellen Beurteilungsunterschiede nach dem Kriterium „Ablegen der Zwischenprüfung“ zu erkennen (vgl. Tabelle 6).

Tabelle 6 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang an Universitäten vor und nach Ablegen einer Zwischenprüfung (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)			
Dimensionen der Studienqualität		Ablegen der Zwischenprüfung	
		nein	ja
	n =	2116	3745
	Untergruppen %	36,1	63,9
Interaktion	Mittelwert	47,2	46,0
	Standardabweichung	19,2	20,5
Didaktik	Mittelwert	52,0	50,3
	Standardabweichung	14,3	15,1
Struktur	Mittelwert	47,5	46,9
	Standardabweichung	14,2	14,8

Quelle: Konstanzer Projekt Studiensituation: Frage 34: „Haben Sie bereits die Zwischenprüfung/das Vordiplom in Ihrem Hauptfach abgelegt?“

Dies gilt auch dann, wenn die Lehrbeurteilung vor und nach dem Ablegen der Zwischenprüfung fachspezifisch betrachtet wird (für einige Beispiele siehe Anhang, Tabelle A1a und A1b). Allerdings sind sehr wohl in manchen Fächern bedeutsame Unterschiede zu erkennen.

Diese Unterschiede sind jedoch nicht systematisch verteilt. In manchen Fächern wird die Lehre in einigen Dimensionen besser beurteilt. Angehende Mathematiker und Mathematikerinnen zum Beispiel bewerten sowohl die Betreuung und Beratung als auch die didaktisch/inhaltliche

Dimension besser. Es gibt Fächer ohne eine Veränderung in der Qualitätseinschätzung im Grund- und Hauptstudium (z.B. Germanistik, Psychologie, Physik oder Maschinenbau). In den Fächern Betriebswirtschaftslehre oder Chemie wird nur die Struktur des Studienfachs kritischer gesehen, in den Rechtswissenschaften die beiden Dimensionen Didaktik und Struktur, und die Mediziner und Medizinerinnen bewerten sowohl die Interaktion mit den Lehrenden als auch die Didaktik schlechter.

Eine systematische Verschiebung in den Urteilen der Studierenden wird damit jedoch nicht deutlich. Daraus läßt sich schließen, daß die Urteile von Studierenden im Grund- oder im Hauptstudium nicht einer Verzerrung durch die Studienphase unterliegen, d.h. das Ablegen der Zwischenprüfung einen unerwünschten Einflußfaktor darstellt. Es ist daher davon auszugehen, daß die Urteilsunterschiede valide sind und reale Veränderungen zwischen Grund- und Hauptstudium nach den Erfahrungen der Studierenden widerspiegeln.

Schlußfolgerung

Die Vermutung, daß Studierende mit „höherer Sachkenntnis“ die Lehre anders beurteilen, kann prinzipiell nicht bestätigt werden. Studierende, die vor Studienaufnahme eine Ausbildung absolviert haben, die fachlich mit dem Studium zusammenhängt, beurteilen die Lehre nicht anders als Studierende, die direkt nach der Schule das Studium aufnehmen.

Gleiches gilt für Studierende, die die Hochschule gewechselt, im Ausland studiert oder an einem Forschungsprojekt mitgearbeitet haben. Ebenfalls scheint sich eine Beschäftigung an der Hochschule als Tutor bzw. Tutorin oder als studentische Hilfskraft nicht auf die Lehrbewertung auszuwirken.

Allerdings ist festzuhalten, daß mit höherer Semesterzahl die Urteile der Studierenden in allen Dimensionen der Lehre kritischer werden. Dies läßt sich aber nicht an den Abschnitten des Grund- und Hauptstudiums festmachen. Auch wenn nicht zu folgern ist, daß die Semesterzahl sich verzerrend auf die studentische Lehrveranstaltungskritik auswirkt, bleibt doch die zunehmend kritischere Bewertung der Studienqualität mit steigender Semesterzahl beachtenswert. Denn sie ist als Hinweis zunehmender Enttäuschung über die Studienbedingungen zu verstehen.

5 Unreife und Unerfahrenheit der Studierenden

Manche Lehrende bezweifeln, daß Studierende die Lehre beurteilen können, da sie zu unreif und unerfahren seien, manche unterstellen sogar, daß ihnen der notwendige Ernst fehle, weshalb ihre Urteile 'beliebig' ausfallen und starke Schwankungen aufweisen würden.

Die Annahme, daß die Unreife oder Unerfahrenheit von Studierenden Einfluß auf ihre Lehrbewertung haben könnten, kann zum einen als vermuteter Mangel an Erfahrung (bzw. fehlende Beobachtungsschulung) gewertet werden, was bereits im vorherigen Abschnitt („Mangelnde Sachkenntnis von Studierenden“) thematisiert wurde. Zum anderen wird dabei auch die Vorstellung von Studierenden als noch-nicht-erwachsen und weniger verantwortungsbewußt thematisiert.

Die Frage nach der Reife und Ernsthaftigkeit der studentischen Beurteiler und Beurteilerinnen wird vor allem dann heftig diskutiert, wenn sich an entsprechende Ergebnisse von Erhebungen wirkliche Konsequenzen (Berufung, Gehalt, etc.) anschließen können. Es gibt einige Untersuchungen, die der Frage nachgegangen sind, inwieweit Studierende ihre Urteile verantwortungsbewußt fällen.

Orpen (1980) ermittelte, ob Studierende sich von der Art der Befragungseinleitung beeinflussen lassen. Kursteilnehmern und -teilnehmerinnen wurde entweder mitgeteilt, daß der Fragebogen nur Feedbackzwecken dienen sollte oder aber, daß er Auswirkungen auf Beförderung, Festanstellung und Gehalt haben würde. Die Ergebnisse zeigten signifikant bessere Bewertungen (1%-Niveau), wenn die Auswertung der Fragebogen administrativen Zwecken dienen sollte. Laura A. Driscoll und Goodwin (1979) kamen bei ihrer Studie zu folgendem Ergebnis: Lehrende erhielten signifikant höhere Werte bei den Instruktionen: 'Administrative Zwecke' und 'Kursverbesserung'.

Aleamoni und Pamela Z. Hexner (1980) schließlich analysierten zwei vergleichbare Gruppen eines Kurses beim gleichen Dozenten. Im ersten Durchlauf wurde der verwendete Fragebogen nur mit einigen Standardformulierungen ausgeteilt, im zweiten Durchlauf wurden die Zwecke der Beurteilung (Bestimmung von Gehaltserhöhung, Beförderung etc.) zusätzlich verbal erläutert. Es konnte festgestellt werden, daß beim zweiten Durchlauf die Bewertung auf allen sechs Subskalen signifikant höher war.

Insgesamt kann nach diesen Studien mit ihren gleichlautenden Befunden höchstens befürchtet werden, daß Studierende bei konsequenzreichen Beurteilungen zu mildereren Urteilen kommen könnten.

Bei unserem Versuch, Ergebnisse zu diesem Themenfeld anhand des Studierendensurveys beizusteuern, sind wir davon ausgegangen, daß 'Erwachsensein' in soziologischen und psychologischen Definitionen meist mit biologischen Faktoren und der Erfüllung bestimmter

kulturspezifischer Lebensaufgaben verbunden wird und die Zeitspanne ab dem dritten Lebensjahrzehnt miteinschließt. Nahezu die gesamte vorliegende Stichprobe der Studierenden ist mindestens 20 Jahre alt und befindet sich daher bereits im dritten Lebensjahrzehnt.

Dennoch wird von seiten der Lehrenden häufig vermutet, daß viele Studierende noch zu unreif und unerwachsen sind, um die Lehre zu beurteilen. In diesem Zusammenhang wird vorgebracht, daß dies gerade für die Studierenden des Grundstudiums gilt. Uns war es nur möglich, einen Indikator für 'Erwachsensein' zu verwenden: das Alter der Studierenden. Diesen Indikator zu verwenden, erscheint jedoch etwas problematisch. Er schließt nämlich zwangsläufig die Annahme ein, daß ältere Menschen grundsätzlich „erwachsener“ wären als jüngere. Wir haben ihn trotzdem benutzt, da ein großer Teil der Bedenken in diesem Bereich sich darauf bezieht. Außerdem ermöglicht er, gemeinsam mit den Befunden des vorigen Abschnitts (mangelnde Sachkenntnis) insgesamt eine Aussage zur Lehrbeurteilung zwischen jüngeren/älteren und erfahreneren/unerfahreneren Studierenden.

Wird innerhalb eines Studienjahrs die Beurteilung Studierender unterschiedlichen Alters betrachtet, kann festgestellt werden, daß es keine durchgängigen Verzerrungen gibt (siehe Anhang, Tabelle A2). Der einzige bedeutsame Unterschied zeigt sich bei Studienanfängern und -anfängerinnen, die älter als 30 Jahre sind: Sie schätzen die Betreuung und Beratung deutlich besser ein. Auch dabei handelt es sich allerdings vermutlich um einen zutreffenden Befund. Denn es ist anzunehmen, daß solche Studierende sich ihre Betreuung und Beratung sowohl anders organisieren können als auch daß Lehrende mit dieser kleinen Gruppe (2-3%) anders umgehen.

Schlußfolgerung

Insgesamt kann festgestellt werden, daß Studierende verantwortungsvoll mit Lehrbeurteilungen umgehen und fehlender Ernst bei der Beantwortung von Fragebogen zu Lehre und Studium nicht erkennbar ist.

Sieht man im Alter einen wichtigen (wenngleich nicht unproblematischen) Indikator für verantwortungsbewußtes Handeln, so kann nicht gefolgert werden, daß die Studierenden ihre Beurteilungen unreif oder unernst treffen würden. Darauf weisen zudem jene Studien hin, die untersucht haben, ob der den Studierenden genannte Verwendungszweck die Einschätzung der Lehrqualität beeinflusst.

Nimmt man zu diesen Ergebnissen noch die Feststellungen des vorigen Abschnitts hinzu, so läßt sich bilanzieren: Wenn Alter und Lebenserfahrung überhaupt einen Rolle spielen, dann zeigen jüngere Studierende höchstens eine allgemein etwas weniger kritische Haltung zur Studien- und Lehrqualität in ihrem Fach, die sich jedoch nicht auf Fragebogen der studentischen Veranstaltungskritik zu einzelnen Lehrveranstaltungen auswirken muß.

6 Das Geschlecht der Studierenden beeinflusst die Bewertung

Die Annahme, daß Studentinnen sowohl Faktoren der Lehrqualität als auch die Hochschule insgesamt anders betrachten als ihre männlichen Kommilitonen, ist auch in Deutschland weit verbreitet. Erwartet wird, daß sich dieser Geschlechtseinfluß auf die Bewertung einzelner Veranstaltungen und Lehrenden niederschlägt. Es gibt einige empirische Befunde zu dieser Vermutung.

Patricia B. Elmore und Pohlmann (1978) ebenso wie Daniel (1995) und Rindermann konnten keinen globalen Geschlechtseinfluß bei den Lehrbewertungen feststellen. Hornbostel und Daniel (1993) ermittelten ebenfalls keine systematischen, fachübergreifenden Urteilsunterschiede zwischen Studentinnen und Studenten. Sie gehen jedoch davon aus, daß es fachspezifisch unterschiedliche Urteilstendenzen gibt: So würden vor allem die Studentinnen der Rechtswissenschaften im Vergleich zu den männlichen Studierenden die Studien- und Lehrqualität kritischer beurteilen.

Ashton (1975) konnte ermitteln, daß zwei Persönlichkeitseigenschaften bei Studentinnen signifikant mit der Lehrbewertung korrelieren, und zwar Konformität (-.54) und Unabhängigkeit (.48). Allerdings korrelierten sie nur mit dem Teil der Lehrbeurteilungen, die auf der 'Beziehungsebene' liegen. Der Autor schließt daraus, daß zwischenmenschliche Werte höher mit der Lehrbewertung von Studentinnen in Beziehung stehen, als daß sie die Bewertung von Studenten beeinflussen.

Die Ergebnisse unserer Auswertungen lassen keine generellen Unterschiede in der Lehrbeurteilung erkennen, die in allen Fächern gleichermaßen anzutreffen wären. *Studentinnen* und Studenten beurteilen in ihrer Gesamtheit die Studienqualität auf den drei verwandten Dimensionen nahezu identisch. Dies gilt auch für jene Dimension der „Interaktion“, die am ehesten Aspekte der „Beziehungsebene“ versammelt: Studentinnen gelangen insgesamt kaum zu anderen Urteilen als die Studenten (vgl. Tabelle 7).

Tabelle 7			
Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach Geschlecht			
(Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)			
Dimensionen der Studienqualität		Geschlecht	
		männlich	weiblich
	n =	5072	3356
	Untergruppen %	60,2	39,8
Interaktion	Mittelwert	47,1	48,4
	Standardabweichung	19,6	20,2
Didaktik	Mittelwert	51,4	51,4
	Standardabweichung	14,8	15,1
Struktur	Mittelwert	47,8	47,3
	Standardabweichung	14,3	14,7

Quelle: Konstanzer Projekt Studiensituation: Frage 120: „Ihr Geschlecht?“

Daß zwischen Männern und Frauen insgesamt in keiner der drei Dimensionen Unterschiede festzustellen sind, könnte an gegenläufigen Tendenzen in den verschiedenen Fächergruppen liegen. Prüft man dies daher für die einzelnen Fächergruppen, so gilt die Aussage ebenso für fast alle Fächergruppen an den Universitäten. Nur im Fach Jura urteilen die Studentinnen etwas positiver im Gegensatz zu den erwähnten Befunden von Hornbostel und Daniel (1993) (vgl. Tabelle 8).

Tabelle 8													
Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach ausgewählten Fächergruppen und Geschlecht an Universitäten													
(Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)													
Dimensionen der Studienqualität		Fächergruppen											
		Kultur		Sozial		Jura		Wirtsch.		Medizin		Natur	
	n =	1278		682		634		872		642		1251	
	Geschlecht	M	W	M	W	M	W	M	W	M	W	M	W
	n =	517	761	241	441	342	292	592	280	312	330	842	409
	Untergruppen %	40,5	59,5	35,3	64,7	53,9	46,1	67,9	32,1	48,6	51,4	67,3	32,7
Interaktion	Mittelwert	56,0	54,8	49,6	51,5	40,3	44,2	36,2	37,3	36,8	35,8	48,7	48,0
	Standardabweichung	18,8	20,4	18,0	17,4	19,3	20,9	17,3	17,5	19,0	18,0	19,0	19,9
Didaktik	Mittelwert	54,0	54,1	49,5	49,9	49,9	51,1	47,3	46,4	50,1	51,2	51,0	51,4
	Standardabweichung	15,2	15,6	14,0	14,3	15,1	16,3	13,9	14,2	14,8	14,3	15,0	15,2
Struktur	Mittelwert	42,8	44,3	43,0	43,8	44,2	47,9	46,0	46,2	49,3	50,4	47,9	48,6
	Standardabweichung	14,1	14,6	14,1	15,0	14,6	14,6	14,9	13,9	14,8	15,0	14,2	14,2

Quelle: Konstanzer Projekt Studiensituation: Fragen 2 (1.Hauptfach) und 120 (Geschlecht)

Ausnahmen sind ebenfalls in den Ingenieurwissenschaften zu finden. Die Aufschlüsselung für einzelne Fächer der Ingenieurwissenschaften an Universitäten und Fachhochschulen läßt keinen einheitlichen Trend in der Lehrbeurteilung nach Geschlecht erkennen. Die Verhältnisse fallen von Fach zu Fach unterschiedlich aus, zumeist sind die Differenzen nicht erheblich (vgl. Tabelle 9).

Tabelle 9 Beurteilung von Interaktion, Didaktik und Struktur in Studiengängen der Ingenieurwissenschaft nach Geschlecht (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)													
Dimensionen der Studienqualität		UNIVERSITÄTEN						FACHHOCHSCHULEN					
		Maschinenbau		Architektur		Bauing.		Maschinenbau		Architektur		Bauing.	
Fach		334		104		196		338		165		183	
n (Fach) = Geschlecht		M	W	M	W	M	W	M	W	M	W	M	W
n =		305	29	63	41	167	29	310	28	103	62	147	36
Untergruppen %		91,3	8,7	60,6	39,4	85,2	14,8	91,7	8,3	62,4	37,6	80,3	19,7
Interaktion	Mittelwert	46,4	47,3	49,7	43,0	42,0	38,5	50,8	58,7	53,1	53,9	50,7	49,5
	Standardabweichung	20,3	20,8	19,1	17,1	16,5	15,9	18,8	15,5	17,8	16,8	19,0	18,1
Didaktik	Mittelwert	51,7	53,3	50,2	45,2	50,0	46,1	54,5	57,4	49,0	53,4	55,1	54,2
	Standardabweichung	14,7	15,7	13,5	14,9	13,1	12,6	15,4	13,3	13,7	15,8	15,2	15,0
Struktur	Mittelwert	51,3	51,6	46,7	46,8	48,8	47,9	50,7	54,4	49,4	54,3	49,2	50,7
	Standardabweichung	13,8	12,4	13,6	14,1	12,9	12,2	13,7	12,5	12,2	14,5	13,3	13,7

Quelle: Konstanzer Projekt Studiensituation: Fragen 124 (Hochschulort/-art), 2 (1. Hauptfach) und 120 (Geschlecht).

Insgesamt ergeben die Daten für Maschinenbau, Architektur und Bauingenieurwesen ein recht inkongruentes Bild. Treten für die Fächer Maschinenbau und Bauingenieurwesen an den Universitäten keine geschlechtsspezifischen Unterschiede auf, so sind diese zumindest im Fach Maschinenbau an den Fachhochschulen sehr wohl zu finden. Frauen beurteilen dort die Beratung und Betreuung signifikant besser als ihre Kommilitonen. Die Studentinnen der Architektur sind an den Universitäten hinsichtlich Betreuung und Beratung dagegen unzufriedener als die Studenten.

Schlußfolgerung

Die Daten des Studierendensurvey geben keine Hinweise darauf, daß Studentinnen die Studienqualität und Lehrsituation insgesamt anders beurteilen als Studenten. Werden fächerspezifisch Unterschiede nach dem Geschlecht gefunden, die vereinzelt vorkommen, kann davon ausgegangen werden, daß sie tatsächlichen Umständen entsprechen; d.h. daß Studentinnen in einigen Fächern tatsächlich andere Gegebenheiten vorfinden oder erleben.

Diese Befunde decken sich daher mit denen von Patricia B. Elmore und Pohlmann (1978), Rindermann (1996) und Hornbostel und Daniel (1993).

7 Unterschiedliches Interesse am Fach, an Studium und Wissenschaft

Immer wieder wird vermutet, daß verschiedene Eigenschaften der Studierenden ihre Lehrbeurteilungen beeinflussen oder gar determinieren. Dabei wird häufig ein unterschiedlich ausgeprägtes Interesse am entsprechenden Fach bzw. Studium als möglicher Faktor genannt.

„Interesse“ wird operationalisiert als Fachinteresse bei der Studienfachwahl, Bedeutung des Studiums, Wichtigkeit von Wissenschaft und Forschung und Anstreben einer wissenschaftlichen Laufbahn, weil dann Fragen und Indikatoren im Studierendensurvey vorliegen. Damit werden drei Facetten des Interesses unterschieden, die sich auf Studium, Fach und Wissenschaft beziehen; sie können sich überschneiden, gehen aber häufig nicht ineinander auf.

Fachinteresse als Studienmotiv

Hofman (1988) prüfte verschiedene Studienmotivationsmerkmale und ihre Auswirkung auf studentische Veranstaltungskritik. Studierende der Psychologie aus verschiedenen Kursen wurden um die Bewertung ihrer Veranstaltungen gebeten. Beurteilt wurden sowohl die einzelnen Sitzungen, die Gesamtbeurteilung, der subjektive Lernfortschritt, das Interesse an verschiedenen psychologischen Themen, Leistungsmotivation, Arbeitsverhalten und Prüfungs- und Sprechängstlichkeit. Das Ergebnis von Korrelationsuntersuchungen zeigte einen positiven Zusammenhang zwischen der Gesamtbewertung der beiden methodischen Veranstaltungen in mittlerer Höhe mit der Einstellung zum methodischen Vorgehen in der Psychologie und dem Interesse an kognitiver und mathematischer Psychologie. Ein negativer Zusammenhang ergab sich mit dem Interesse an psychoanalytischer und Sexualpsychologie. Korreliert wurden auch einzelne Dimensionen mit den Bewertungen. Zwischen 22 und 47 Prozent der Varianz der mittleren Beurteilungen einzelner Dimensionen ließen sich durch Kombinationen von zwei oder drei Motivationsmerkmalen vorhersagen. Insgesamt schließt Hofmann: "Die Hypothese, daß Merkmale der Studienmotivation eine tendenzielle Determinante der Beurteilung von Veranstaltungen darstellen, wird somit insgesamt gut bestätigt" (Hofman, 1988, S. 125).

Kromrey (1993, 1994, 1995) geht sogar so weit, daß er die unterschiedlichen Lehrbewertungen vor allem auf das unterschiedliche Interesse der Studierenden am entsprechenden Kurs zurückführt. Auch Rindermann (1996) fand beträchtliche Übereinstimmungen ($r=.46$) zwischen dem Vorinteresse der Studierenden am Kursthema und ihrer anschließenden Beurteilung der Lehrqualität des entsprechenden Kurses. Daniel (1994) konnte ergänzend ermitteln, daß weniger interessierte Studierende die Lehrenden zwar etwas schlechter bewerten, daß jedoch die Rangfolge der Bewertung mit denen interessierter Studierender identisch ist und sich nur parallel verschiebt.

Ähnliche Unterschiede sind ebenfalls zu erwarten, wenn die Beurteilungen von Pflicht- oder Wahlkursen verglichen werden. Denn bei Wahlkursen ist davon auszugehen, daß Studierende stärker aufgrund eigenen Interesses die Veranstaltung besuchen als aufgrund studienorganisatorischer Zwänge.

Brandenburg, Slinde und Batista (1977) fanden entsprechende Zusammenhänge zwischen der Lehrbewertung und der Variable Pflicht- oder Wahlfach. Auch in der Untersuchung von Marcia K. Petchers und Chow (1988) wurde der Frage nachgegangen, inwieweit externe Rahmenbedingungen Einfluß auf studentische Lehrbewertungen haben. Einzig der Faktor 'Wahl- oder Pflichtkurs' beeinflusste zwei der drei verwendeten Skalen signifikant.

Eine mögliche Erklärung dafür liefern DuCette und Jane Kenney (1982), die herausfanden, daß in bestimmten Kursen stärker als in anderen die Lehrbewertung mit den Noten korreliert (z.B. in Statistikkursen). Dazu vermuten DuCette und Jane Kenney, daß in dieser Art von Kursen die Studierenden kein großes inhaltliches Interesse haben und Noten daher um so wichtiger werden.

Eine andere Erklärungsmöglichkeit wäre, daß Pflichtkurse häufig methodische oder sonstige, eher als unangenehm empfundene Kurse (z.B. Einführung in die Statistik) sind. Wie gut sie gemeistert werden, hängt dabei stärker als in anderen Kursen von Vorkenntnissen und Neigungen ab, die sich wiederum auf die Lernmotivation auswirken. Schätzen Studierende z.B. ihre erwartete Statistiknote gut ein, so kann man davon ausgehen, daß sie im entsprechenden Kurs gut mitgekommen sind.

Die aufgeführten Studien beziehen sich auf die Übereinstimmung zwischen dem Interesse der Studierenden an den Themen einzelner Kurse und ihrer Beurteilung der Veranstaltungen. Wir haben im Unterschied dazu überprüft, ob sich ihr allgemeines *Fachinteresse* auf die allgemeine Lehreinschätzung auswirkt (vgl. Tabelle 10).

Insgesamt ist zu vermuten: Je wichtiger das Fachinteresse bei der Studienfachwahl war, desto besser wird die Lehre beurteilt. Im statistischen Sinne zeigen sich zwar signifikante Unterschiede auf den beiden Dimensionen 'Interaktion' und 'Didaktik'; jedoch nur zwischen der verschwindend kleinen Minderheit von 2,1 Prozent der Studierenden, denen das Interesse am Fach bei der Studienwahl weniger wichtig war und der breiten Mehrheit (69,6%), denen dies sehr wichtig war. Die Unterschiede sind bei der Beurteilung der Struktur des Studienganges auch vorhanden, bleiben aber knapp unter der Grenze von fünf Prozentpunkten (als Maß für bedeutsame Differenzen).

Tabelle 10 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach Wichtigkeit des speziellen Fachinteresses bei der Studienfachwahl (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)				
Dimensionen der Studienqualität		Spezielles Fachinteresse war für die Studienfachwahl ... (Skala von 0 (= unwichtig) bis 6 (sehr wichtig))		
		wenig wichtig (0 - 1)	teilweise wichtig (2 - 4)	sehr wichtig (5 - 6)
	n =	175	2381	5846
	Untergruppen %	2,1	28,3	69,6
Interaktion	Mittelwert	40,5	44,6	49,1
	Standardabweichung	20,5	18,6	20,2
Didaktik	Mittelwert	46,0	49,6	52,3
	Standardabweichung	15,5	14,3	15,1
Struktur	Mittelwert	43,7	46,0	48,3
	Standardabweichung	15,0	14,1	14,5

Quelle: Konstanzer Projekt Studiensituation: Frage 14.1: „Wie wichtig waren Ihnen die folgenden Gründe bei der Entscheidung für Ihr jetziges Studienfach: spezielles Fachinteresse?“

Bedeutung des Studiums

Mittlerweile ist ein großer Teil der Studierenden nebenbei oder hauptsächlich berufstätig. Für immer weniger von ihnen ist das Studium über die Anfangssemester hinaus zentraler Mittelpunkt in ihrem Leben. Das schlägt sich auch in den Diskussionen um ‘Proforma’- und ‘Teilzeitstudierende’ nieder, wobei unterstellt wird, daß sich die geänderte Einstellung zum Lebensabschnitt ‘Studium’ auf die Beurteilung der Lehrqualität auswirkt.

Vorab ist zu bemerken, daß nur ein gutes Viertel der Studierenden Hochschule und Studium wirklich als ihren Lebensmittelpunkt betrachtet. Für die meisten Studierenden sind andere Bereiche mindestens ebenso wichtig. Für viele ist der Besuch einer Hochschule ein Zeitraum des frühen Erwachsenenalters, in dem auch gearbeitet wird, andere Möglichkeiten der Fortbildung genutzt und in manchen Fällen sogar Familien gegründet werden. Diese Tendenz schlägt sich im großen und ganzen aber nicht in der Beurteilung der Lehrqualität nieder (vgl. Tabelle 11).

Allerdings gibt es eine Ausnahme: die fünf Prozent der Studierenden, die in ihrem Studium eher eine Nebensache sehen, sind mit der Lehre in allen Bereichen weniger zufrieden. Aufgrund der vorliegenden Daten ist es jedoch nicht möglich, kausale Zusammenhänge zu bestimmen. Es kann somit nicht gesagt werden, ob eine größere Unzufriedenheit mit den Bedingungen an den Hochschulen eher zu einer Entfernung von Studium und Hochschule und verstärkten Nebentätigkeiten führt oder umgekehrt. Zu vermuten ist allerdings, daß bei dem beschriebenen Phänomen eine sich wechselseitig verstärkende Beziehung vorliegt.

Tabelle 11 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach der Bedeutung des Studiums für die derzeitige Lebenssituation der Studierenden (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)					
Dimensionen der Studienqualität		Bedeutung des Studiums			
		Studium als Lebensmittel- punkt	Studium ist wie Berufstätigkeit	Andere Berei- che sind gleich wichtig	Studium ist eher Nebensa- che
n =		2340	3059	2367	413
Untergruppen %		28,6	37,4	28,9	5,0
Interaktion	Mittelwert	49,2	47,6	47,1	42,08
	Standardabweichung	20,7	19,2	19,6	20,3
Didaktik	Mittelwert	53,3	52,0	50,0	45,54
	Standardabweichung	15,3	14,4	14,7	16,13
Struktur	Mittelwert	48,8	48,7	46,0	41,95
	Standardabweichung	14,7	13,9	14,5	14,62

Quelle: Konstanzer Projekt Studiensituation: Frage 77: „Welche der folgenden Beschreibungen würde Ihre derzeitige Lebenssituation am ehesten kennzeichnen?“

Relevanz von Wissenschaft und Forschung

Neben der eher den studentischen Lebensraum umfassenden Frage nach der Bedeutung des Studiums, stellt sich zusätzlich die Frage nach der stärker inhaltlichen Konnotation des Studiums, nämlich der Relevanz von Wissenschaft und Forschung.

Je wichtiger Studierende Wissenschaft und Forschung nehmen, desto besser beurteilen sie die Lehre (vgl. Tabelle 12). Zwischen den Studierenden, die Wissenschaft und Forschung für sich als unwichtig einstufen, und denen, die sie als wichtig betrachten, sind statistisch bedeutsame Unterschiede in den Dimensionen Interaktion und Didaktik festzustellen.

Studierende mit höherem Interesse an Wissenschaft und Forschung haben häufiger fachliche Kontakte innerhalb der Hochschule (vor allem in den Univesitätsstudiengängen) und ein größeres Interesse an den Studieninhalten. Vermutlich haben sie einen höheren Lerngewinn im Studium und bewerten deshalb die Studienqualität folgerichtig besser.

In diesem Zusammenhang stellt sich die Frage nach dem *potentiellen wissenschaftlichen Nachwuchs* und seiner Einstellung zu Lehre. Interessant ist dieser Punkt vor allem, weil aus dieser Gruppe sich die nächste Generation der Lehrenden rekrutiert, die die Studien- und Lehrqualität maßgeblich mitbestimmen wird. Beurteilen somit angehende Wissenschaftler und Wissenschaftlerinnen die Lehre anders, möglicherweise besser, und wenn dies der Fall sein sollte, was sind die auslösenden Gründe?

Als wissenschaftlicher Nachwuchs wird der von Gawatz (1991) beschriebene ‘Wissenschaftlertypus’ herangezogen. Dieser Studierendentypus hat bereits während der Studienzeite eine

stark wissenschaftliche Ausrichtung, betrachtet Wissenschaft und Forschung als wichtigen Lebensbereich und will später wissenschaftlich tätig sein.

Tabelle 12 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang in Abhängigkeit der Wichtigkeit von Wissenschaft und Forschung für die Befragten (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)				
Dimensionen der Studienqualität		Bedeutung von Wissenschaft und Forschung (Skala von 0 (= völlig unwichtig) bis 6 (sehr wichtig))		
		unwichtig (0 - 2)	teilweise wichtig (3 - 4)	wichtig (5 - 6)
	n =	2505	3969	1926
	Untergruppen %	29,8	47,3	22,9
Interaktion	Mittelwert	44,6	48,0	51,0
	Standardabweichung	19,5	19,5	20,5
Didaktik	Mittelwert	48,9	51,8	53,9
	Standardabweichung	14,9	14,3	15,7
Struktur	Mittelwert	45,9	48,0	48,7
	Standardabweichung	14,7	13,9	15,1

Quelle: Konstanzer Projekt Studiensituation: Frage 101.5: „Geben Sie bitte an, wie wichtig die einzelnen Lebensbereiche für Sie persönlich sind : Wissenschaft und Forschung?“

Studenten und Studentinnen dieses Typus, es sind übrigens nahezu gleich viele Männer wie Frauen, sind etwas leistungsorientierter. Sie können zu 25,6 Prozent zum Typ des 'Maximalisten' hinsichtlich Effizienz und Studienerfolg gezählt werden, der ansonsten nur zu 14,6 Prozent vertreten ist (vgl. dazu auch den folgenden Abschnitt 8). Solche Studierenden sind bereit, schnell, intensiv und erfolgsorientiert zu studieren. Sie haben überdurchschnittlich oft eine Stelle als studentische Hilfskraft (35,7 % im Vergleich zu 15,2%) und, falls sie nicht bereits eine Stelle haben, so doch ein höheres Interesse an ihr (40,8%) als andere Studierende (33,1%). Dies erscheint plausibel, da solche Stellen immer noch als Ausgangspunkt für eine wissenschaftliche Laufbahn betrachtet werden können.

Tatsächlich bewertet der „wissenschaftliche Nachwuchs“ unter den Studierenden die Lehre besser als Studierende mit anderen beruflichen Perspektiven (vgl. Tabelle 13).

Dies gilt jedoch statistisch bedeutsam nur für die Betreuung und Beratung durch die Lehrenden. Zu vermuten ist, daß dies vor allem durch die stärkere Bedeutung der studentischen Hilfskraftstellen und die etwas ausgeprägtere Leistungsorientierung zu erklären ist.

Tabelle 13

Beurteilung von Interaktion, Didaktik und Struktur im Studiengang in Abhängigkeit mit der Übereinstimmung mit dem „Wissenschaftlertypus“ nach GAWATZ

(Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)

Dimensionen der Studienqualität		Wissenschaftlertypus	
		nein	ja
	n =	7779	682
	Untergruppen %	91,9	8,1
Interaktion	Mittelwert	47,1	53,5
	Standardabweichung	19,8	19,9
Didaktik	Mittelwert	51,0	55,9
	Standardabweichung	14,8	16,1
Struktur	Mittelwert	47,3	50,2
	Standardabweichung	14,4	14,5

Quelle: Konstanzer Projekt Studiensituation: Fragen 17.5, 17.6, 93.10, 93.11 und 101.4.

Schlußfolgerung

Die Ergebnisse deuten darauf hin, daß je stärker das Fachinteresse als Motiv der Studienwahl ausschlaggebend war, desto besser werden Dimensionen der Lehrqualität eingeschätzt. Dies läßt die Vermutung zu, daß ein höheres Fachinteresse insgesamt zu einer besseren Einschätzung der erlebten Studien- und Lehrqualität führt.

Ob aber das Studium als zentraler Lebensmittelpunkt betrachtet wird, hat keine nennenswerten Auswirkungen auf die Lehrbeurteilung.

Studierende, denen Wissenschaft und Forschung wichtiger ist als anderen, bewerten die Studienqualität insgesamt etwas besser. Diejenigen unter ihnen, die auch eine wissenschaftlichen Berufsperspektive haben, beurteilen besonders die Betreuung und Beratung durch die Lehrenden besser als andere Studenten und Studentinnen.

Dafür bieten sich verschiedene Erklärungsmöglichkeiten an: diese Studierende haben überdurchschnittlich häufig hilfswissenschaftliche Stellen und schon allein dadurch mehr Kontakt zu Lehrenden. Sie sind höher motiviert, das Studium schnell und vor allem gut abzuschließen, und sie identifizieren sich unter Umständen stärker mit der Rolle des Lehrenden und beurteilen daher die Dozenten und Dozentinnen etwas besser.

Anzunehmen ist, daß Studierende mit einem höheren Fachinteresse motivierter und leistungsbereiter sind, bereits über mehr Kenntnisse in ihrer Disziplin verfügen, sich neue Themen leichter aneignen und mehr Erfolgserlebnisse haben. Dies wirkt sich dann auch auf ihre Lehrbewertungen aus.

Die Befunde anhand des Studierendensurveys decken sich mit Untersuchungsergebnissen anderer Studien. Es erscheint daher sinnvoll, bei Studien- und Lehrbewertungen, sei es ganzer Fachbereiche oder einzelner Veranstaltungen, das jeweilige Fach- bzw. Kursinteresse mitzuerheben, um keine unerwünschten Verzerrungen bei Vergleichen zu riskieren. Auswertungen sollten daher das vorhandene Interesse der Studierenden berücksichtigen und die Daten auch nach dem Grad des studentischen Interesses ausweisen.

8 Unterschiedliche Leistungsbereitschaft

Unterschiedliche Leistungsbereitschaft wird meist unter dem Begriff der Leistungsmotivation mit anderen Eigenschaften der Studierenden diskutiert, die in Verdacht stehen, Lehrbeurteilungen zu verzerren. Unter Leistungsmotivation wird die Tendenz verstanden, die eigene Leistungsfähigkeit in verschiedenen Bereichen unter Beweis zu stellen und zu steigern. Angenommen wird, daß unterschiedlich leistungsmotivierte bzw. leistungsbereite Studierende auch die Lehre anders beurteilen.

Kovac (1976) ermittelte, daß Studierende mit einem hohen Leistungsanspruch Inhalt und Struktur eines Kurses besser als Studierende mit geringerem Leistungsanspruch bewerteten. Strom et.al. (1982) konnten feststellen, daß leistungsmotivierte Studierende weniger strukturierte und anspruchsvollere Kurse vorziehen.

In der Studie von Pamela J. Cooper et. al. (1982) wurden im Vergleich zu der vorherigen Untersuchung keine Zusammenhänge zwischen der Lehrbewertung und der Leistungsmotivation der Studierenden ermittelt. Problematisch dabei ist eine sehr niedrige Rücklaufquote pro Kurs, die u.U. zu dem diskrepanten Ergebnis geführt haben könnte.

Der verwendete Datensatz des Studierendensurveys bietet keine Möglichkeiten, die Leistungsmotivation der Studierenden im Sinne eines psychometrisch erfassbaren Merkmals (etwa durch Tests) zu ermitteln. Es gibt jedoch einige Items, die Hinweise auf das Ausmaß der Leistungs- und Anstrengungsbereitschaft der Studierenden geben.

Ein möglicher Indikator wäre die Bereitschaft des Studenten, der Studentin *intensiv zu Arbeiten, um ein gutes Examen zu erreichen*. Dieses Item wurde von den Befragten auf einer Skala von 0 (trifft überhaupt nicht zu) bis 6 (trifft voll und ganz zu) eingestuft.

Tabelle 14

Beurteilung von Interaktion, Didaktik und Struktur im Studiengang in Abhängigkeit des Leistungsanspruchs der Studierenden

(Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)

Dimensionen der Studienqualität		Intensives Arbeiten für ein gutes Examen (Skala von 0 = trifft überhaupt nicht zu bis 6 = trifft voll und ganz zu)		
		nicht zutreffend (0 - 1)	teilweise zutreffend (2 - 4)	zutreffend (5 - 6)
	n =	857	5313	2246
	Untergruppen %	10,2	63,1	26,7
Interaktion	Mittelwert	44,7	47,5	49,0
	Standardabweichung	19,6	19,3	21,1
Didaktik	Mittelwert	47,9	51,2	53,1
	Standardabweichung	14,9	14,5	15,7
Struktur	Mittelwert	43,6	47,3	49,6
	Standardabweichung	14,4	14,1	15,0

Quelle: Konstanzer Projekt Studiensituation: Frage 25.2: „In wieweit treffen folgende Aussagen über Lernen und Studieren auf Sie persönlich zu: Ich arbeite sehr intensiv, um ein gutes Examen zu erreichen?“

Studierende, die in einem höheren Maß bereit sind, für ein gutes Examen intensiv zu arbeiten, unterscheiden sich in bezug auf ihre Lehrbewertung zwar tendenziell von den Studierenden, die dazu weniger bereit sind, die Differenzen sind aber nicht so groß, daß sie als bedeutsam einzustufen wären (vgl. Tabelle 14).

Nimmt man zum beabsichtigten Arbeitsaufwand noch den Vorsatz hinzu, das Studium möglichst rasch abzuschließen, lassen sich sieben Studiertypen bilden, die sich hinsichtlich ihrer „subjektiven Studieneffizienz“ unterscheiden (vgl. Bettina Leitow, 1996).

Bettina Leitow unterscheidet nach Analyse verschiedener Studierhaltungen und Studienstrategien drei Studiertypen, die hauptsächlich durch ihren erfolgsorientierten Ehrgeiz bestimmt sind: *Minimalisten*, die eine Effizienzorientierung in ihrem Studium ablehnen und kaum ehrgeizig sind; *Schlenderer*, die bereit sind, Leistungen auf mittlerem Niveau zu erbringen, ungeachtet ihrer Studiendauer, und *Motivierte*, denen die Studiendauer relativ weniger wichtig ist, die aber sehr intensiv und erfolgsbezogen studieren.

Außerdem werden zwei Typen von Studierenden beschrieben, die ihr Interesse stärker auf einen schnellen Studienabschluß als auf einen guten Studienabschluß legen. Dies sind zum einen die *Sprinter*, die fast ausschließlich Wert auf ein kurzes Studium und fast keinen Wert auf ein gutes Examen legen; zum anderen die *Ambitionierten*, denen ein schneller Abschluß zwar wichtiger ist als ein gutes Examen, dem sie aber doch einige Bedeutung beimessen.

Das Mittelfeld der Studierenden, die sog. *Mittleren Studierenden*, streben gleichermaßen ein mehr oder weniger guten wie schnellen Studiumabschluß an. Bleibt noch die Gruppe der *Ma-*

ximalisten, die sich für ein ebenso schnelles wie erfolgreiches Studium engagieren und insgesamt als besonders ehrgeizig erscheinen.

Diese Studiertypen sind nicht über alle Fächer gleich verteilt. Minimalisten sind am stärksten in den Sozialwissenschaften anzutreffen (sowohl Uni als auch FH) und kaum unter den Juristen und Juristinnen oder den Studierenden der Medizin zu finden. In diesen Fächern sind die Maximalisten überdurchschnittlich vertreten. Insgesamt kann festgestellt werden, daß der Anteil gering bzw. hoch leistungsmotivierter Studierender über die Fachgruppen hinweg unterschiedlich verteilt ist.

Betrachtet man sich die durchschnittliche Lehrbeurteilung der verschiedenen Studiertypen, so wird erkennbar: Je eher Studierende bereit sind, viel für das Examen zu leisten, und je schneller sie ihr Studium abschließen wollen, desto besser beurteilen sie die Studienqualität (vgl. Tabelle 15).

Tabelle 15								
Beurteilung von Interaktion, Didaktik und Struktur nach Studiertypen								
(Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)								
Dimensionen der Studienqualität		Studiertypen						
		Minimalisten	Sprinter	Schleuderer	Mittlere Studierende	Ambitionierte	Motivierte	Maximalisten
n =		934	1184	1011	1400	1621	950	1288
Untergruppen %		11,1	14,1	12,1	16,7	19,3	11,3	15,4
Interaktion	Mittelwert	45,3	46,3	46,8	46,3	49,5	47,9	49,8
	Standardabweichung	19,1	19,6	19,2	18,5	19,9	21,0	21,1
Didaktik	Mittelwert	48,2	49,8	50,2	51,3	52,7	51,7	54,2
	Standardabweichung	14,4	14,9	14,1	14,0	14,9	15,7	15,6
Struktur	Mittelwert	44,0	46,0	45,4	47,8	48,9	48,1	50,7
	Standardabweichung	14,1	14,4	14,2	13,2	14,5	15,0	14,9

Quelle: Konstanzer Projekt Studiensituation Fragen 25.2 und 25.7: „In wie weit treffen folgende Aussagen über Lernen und Studieren auf Sie persönlich zu: Ich arbeite sehr intensiv, um ein gutes Examen zu erreichen? / Mir kommt es darauf an, mein Studium möglichst rasch abzuschließen?“

Diese bessere Studienbeurteilung schlägt sich statistisch bedeutsam jedoch nur in den Dimensionen Didaktik und Struktur nieder. Eine höhere Leistungsbereitschaft hat außerdem Einfluß auf den Studienerfolg: Maximalisten erreichen im Vergleich zu den anderen Gruppen bessere, Minimalisten schlechtere Vordiploms- bzw. Zwischenprüfungsnoten. Diese Unterschiede erweisen sich allerdings nur in einigen Fächern als bedeutsam, vor allem in der Medizin.

Schlußfolgerung

Die beschriebenen Studiertypen zwischen „Minimalisten“ und „Maximalisten“ können durchaus als Ausdruck unterschiedlicher Leistungshaltungen gelten. Die Daten lassen sich

als Hinweis darauf deuten, daß hochmotivierte Studierende die Lehre besser beurteilen als weniger motivierte.

Studierende, die in einem höheren Maße bereit sind, für ein gutes Examen zu arbeiten, unterscheiden sich in bezug auf ihre Lehrbewertung allerdings nicht von den Studierenden, die dies in einem geringeren Maße sind. Anzunehmen ist daher, daß mit dieser Variable eher unterschiedlich ausgeprägter Fleiß als unterschiedliche Leistungsorientierungen erfaßt werden.

Die Ergebnisse unserer Auswertungen ergeben, daß höher leistungsmotivierte Studierende, die viel für ihr Examen leisten und einen schnellen Studienabschluß anstreben, die Studienqualität zumindest in ihrer didaktisch-tutorialen Dimension und hinsichtlich der Struktur ihres Studienganges besser beurteilen als ihre Kommilitonen und Kommilitoninnen. Diese Schlußfolgerung deckt sich mit den Ergebnissen aus den aufgeführten anderen Studien.

9 Noten und Leistungsfähigkeit

Nach einer Untersuchung von Marsh (1982b) befürchten zwei Drittel aller Lehrenden, daß sich ihre Notengebung auf die studentische Lehrbeurteilung auswirkt. Es ist somit nicht verwunderlich, daß Noten eine der meistuntersuchten möglichen Einflußvariablen der studentischen Veranstaltungskritik sind.

Auch Leistungsfähigkeit wird meist in Zusammenhang mit den Noten thematisiert. Unter diesen beiden Stichworten findet sich ein undifferenziertes Konglomerat verschiedener, nicht präzise voneinander abgegrenzter Annahmen, die sich auf ganz unterschiedliche Ebenen möglicher unerwünschter Verzerrungen in der studentischen Veranstaltungskritik richten. Vermutet wird sowohl, daß milder bewertende Lehrende ihrerseits bessere Beurteilungen erhalten (i.S. von „Bestechung“), daß Studierende, die sich ungerecht beurteilt fühlen, deswegen schlechtere Noten für die Lehre vergeben (i.S. von „Rache“) und daß die unterschiedliche Tüchtigkeit der Studierenden sich auf ihre Lehrbewertung auswirkt (i.S. von Unterschieden in der Leistungsfähigkeit).

Es gibt zwar eine beträchtliche Anzahl an Studien in diesem Bereich, die meisten sind jedoch mit Mängeln behaftet. Häufig werden zum Beispiel nicht tatsächlich erhalten, sondern nur erwartete Noten in die Untersuchung einbezogen. Eine Korrelation: erwartete schlechte Note - schlechte Lehrbewertung kann aber nicht die Frage nach einer möglichen „Bestechlichkeit“

oder „Rache“ klären. Vielmehr kann davon ausgegangen werden, daß der Student wenig verstanden hat und daraus schließt, daß der Lehrende den Stoff schlecht vermitteln kann.

Dies würde sich auch mit einer Untersuchung von Snyder und Clair (1976) decken, die ermittelten, daß je besser die Note ausfiel, desto eher wurde internal attribuiert (eigene Leistung), je schlechter die Note war, desto eher kam es zu externalen Attributionen (äußere Faktoren). Snyder und Clair (1976) fanden weiter heraus, daß Studierende, die besser als sie erwarteten beurteilt wurden, ihrerseits bessere Bewertungen abgaben. Am schlechtesten fiel die Lehrbewertung dann aus, wenn die erhaltene Note die Erwartungen enttäuscht hatte.

Yates und Ksarmos (1971) versuchten zu überprüfen, ob Studierenden sich wirklich für schlechte Noten 'rächen' würden. Studierende eines Kurses beantworteten zweimal innerhalb kürzester Zeit (ca. 1 Woche) einen Fragebogen zur Lehrqualität. Der Zweck der zweimaligen Bewertung war ihnen nicht bekannt. Die beiden Wissenschaftler fanden keine Unterschiede in der Art der Bewertung von gut vs. schlecht benoteten Studierenden. Kritisch ist zum einen der kurze Zeitraum zwischen den Erhebungen, zum anderen, daß den Studierenden bei diesem Quasiexperiment kein glaubhafter Grund für den zweiten Fragebogen genannt wurde. Diese Studie ist weniger wegen ihres Ergebnisses interessant, sondern vielmehr als Beispiel einer Anordnung für eine Untersuchung. Auf diese Art und Weise könnte man methodisch genauer dem vermuteten Zusammenhang zwischen Noten und Bewertung nachgehen.

Insgesamt kann noch als ungeklärt gelten, ob sich Studierende durch eine schlechtere Beurteilung für eine enttäuschende Note rächen wollen. Möglich ist z.B. auch, daß dies nur für einige Studierende zutrifft. So verweist z.B. Blass (1980) auf Ergebnisse, die belegen, daß relativ subjektiv bewertende Studierende eine weitaus höhere Korrelation zwischen Note und Lehrbewertung aufweisen als die Gesamtgruppe.

Genau wie bei der Frage, ob sich bei Studierenden schlechte Benotungen tatsächlich in ihrer Lehrbeurteilung niederschlagen, ist der Verdacht möglicher Bestechlichkeit - milde beurteilende Lehrende bekommen ihrerseits gute Bewertungen von den Studierenden geschenkt- im Grunde nicht untersucht. Zwar sind unter diesem Schlagwort eine ganze Reihe von Studien zu finden, sie alle beinhalten jedoch so viele mögliche Faktoren, daß kausale Schlußfolgerungen nicht gezogen werden können.

Gefundene Korrelationen zwischen der Durchschnittsnote in verschiedenen Kursen und der Beurteilung der jeweiligen Lehrenden (z.B. Brown, 1976, Centra, 1977) können verschiedene Gründe haben. Es kann am Kursthema liegen, oder ein didaktisch versierter Lehrender kann einen höheren Lernfortschritt ermöglichen, daher zu recht 'bessere' Studierende mit einem

besseren Notendurchschnitt haben und folglich auch selbst eine bessere Beurteilung erhalten. Eine eindeutige kausale Zuordnung ist aufgrund der bisherigen Studien jedoch nicht möglich.

Wir haben überprüft, ob Studierende, die sich in ihrer *Leistung falsch (besser oder schlechter) eingeschätzt* fühlen, die Lehre anders beurteilen als diejenigen Studierenden, die glauben, „gerecht“ benotet worden zu sein (vgl. Tabelle 16).

Tabelle 16 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang in Abhängigkeit der Übereinstimmung von persönlicher Leistungseinschätzung und erhaltenen Noten (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)				
Dimensionen der Studienqualität		Erhaltene Noten sind ...		
		viel / etwas schlechter	gleich	etwas / viel besser
n =		2505	3969	1926
Untergruppen %		29,8	47,3	22,9
Interaktion	Mittelwert	44,6	48,0	51,0
	Standardabweichung	19,5	19,5	20,5
Didaktik	Mittelwert	48,9	51,8	53,9
	Standardabweichung	14,9	14,3	15,7
Struktur	Mittelwert	45,9	48,0	48,7
	Standardabweichung	14,7	13,9	15,1

Quelle: Konstanzer Projekt Studiensituation: Frage 36.1: „Wenn Sie Ihre bisherigen Studienleistungen insgesamt betrachten, inwieweit stimmen die erhaltenen Noten mit Ihren eigenen Leistungseinschätzungen überein?“

Studierende, deren eigene Leistungseinschätzung von ihren Noten abweicht (dies ist immerhin bei mehr als der Hälfte der Studierenden der Fall), beurteilen die Lehre nur wenig anders als diejenigen, die das Gefühl haben, korrekt beurteilt worden zu sein.

Jedoch wirkt sich das Gefühl der Leistungsunterschätzung bzw. -überschätzung bei der Beurteilung der Betreuung und Beratung (Interaktion) und der Einschätzung der inhaltlich-didaktischen Dimension (Didaktik) aus. Dabei unterscheiden sich statistisch bedeutsam die beiden Gruppen voneinander, die das Gefühl haben, zu gut oder zu schlecht beurteilt worden zu sein. Mit einer besseren Beurteilung als erwartet geht eine bessere Lehrbeurteilung einher, mit einer enttäuschenden Benotung eine schlechtere Bewertung der Lehrqualität.

Problematisch ist allerdings, daß diese Ergebnisse mit mindestens einem Faktoren konfundiert sind: Studierende, die sich unterschätzt fühlen, gehören eher zu den schlecht benoteten, die sich überschätzt fühlen eher zu den gut benoteten. Anzunehmen ist daher, daß sie sich auch in ihrer Leistungsstärke unterscheiden, ein Faktor, der sich auf die Lehrbewertung auswirkt, wie im folgenden Abschnitt dokumentiert wird.

Ob milde bewertende Lehrende bessere Beurteilungen erhalten als streng bewertende, kann nach diesen Ergebnissen nicht eindeutig gesagt werden. Um dies zu klären, wäre eine quasi-experimentelle Anordnung notwendig.

Leistungsstärkere und leistungsschwächere Studierende

Ähnlich wie bei dem Faktor ‘Interesse’ erscheint es plausibel, daß leistungsstärkere Studierende möglicherweise über mehr Potentiale verfügen und auch mehr Zeit in ihr Studium investieren, weshalb Sie dem Stoff besser folgen können. Möglicherweise kommen sie mit dem Studium insgesamt besser zurecht, haben mehr Erfolgserlebnisse und beurteilen daher die Lehre insgesamt günstiger.

Leistungsstärke wird in den meisten Untersuchungen als Notenvariable operationalisiert. Die Daten von Studien in diesem Bereich sind insgesamt diskrepant.

Snyder und Clair (1976) und Brandenburg, Slinde und Batista (1977) fanden signifikant positive Korrelationen zwischen der Lehrbewertung und den erwarteten Noten. Aleamoni und Hexner (1980) und Lester (1982) konnten dagegen keine nennenswerte Korrelation zwischen den Ergebnissen der Lehrbewertung und den späteren Noten ermitteln.

Wir haben überprüft, ob die Vordiploms- bzw. Zwischenprüfungsnote der Studierenden mit ihrer Beurteilung der Lehrqualität zusammenhängt. Festzustellen ist, daß zwischen Note und der Lehrbewertung keine prinzipiellen Zusammenhänge bestehen. Allerdings korreliert die Vordiplomsnote um $r=.20$ mit der Beurteilung der Betreuungs- und Beratungsqualität (vgl. Tabelle 17).

Tabelle 17 Korrelationstableau: Durchschnittsnote der Zwischenprüfung und Qualität von Interaktion, Didaktik und Struktur				
Variable	Skalierung	Mittelwert	Standardabweichung	Besetzungszahl (N)
1. Note Zwischenpr.	1,0 bis 6,0	2,6	0,7	7119
2. Interaktion	1 (= sehr schlecht) bis 100 (= sehr gut)	47,6	19,9	8455
3. Didaktik	1 (= sehr schlecht) bis 100 (= sehr gut)	51,4	14,9	8444
4. Struktur	1 (= sehr schlecht) bis 100 (= sehr gut)	47,6	14,5	8457
N = 8457	1. Note¹⁾	2. Interaktion	3. Didaktik	4. Struktur
1. Note	1,00			
2. Interaktion	- 0,20	1,00		
3. Didaktik	- 0,09	0,62	1,00	
4. Struktur	- 0,03	0,34	0,46	1,00

Quelle: Konstanzer Projekt Studiensituation: Frage 35: „Geben Sie bitte auch die Durchschnittsnote Ihrer Zwischenprüfung/des Vordiploms an.“

1) Die negativen Vorzeichen erklären sich dadurch, daß in Deutschland höhere Noten numerisch niedriger ausfallen. Je besser die Lehrbewertung, desto ‘kleiner’ ist die Note. Daher die negative Korrelation.

Auch hier können fachspezifische Verzerrungen vorliegen. Aufgrund der vorliegenden Daten bietet sich eine fachspezifische Gegenüberstellungen vor allem in bezug auf unterschiedliche Vordiplomsnoten an (vgl. Anhang 2, Tabelle A4). Die Ergebnisse der exemplarisch ausgewählten Studiengänge lassen die Vermutung zu, daß leistungsschwächere Studierende die Lehre tendenziell schlechter beurteilen. Etwas deutlicher werden die Unterschiede vor allem in lern- und arbeitsintensiven Studiengängen, am deutlichsten in der Medizin.

Leistungsschwächere Studierende beurteilen die Lehre nicht insgesamt schlechter, sondern sie stehen gerade besonders problematischen Bereichen ihres Faches bezüglich der Lehr- und Studienqualität weitaus kritischer gegenüber. So beurteilen zum Beispiel die schlechter benoteten Studierenden der Fächer Medizin und Maschinenbau vorwiegend die Betreuung und Beratung schlechter, die Studierenden der Chemie den Aufbau ihres Studiengangs und die Studierenden der Rechtswissenschaften haben eine etwas kritischere Haltung gegenüber der inhaltlich-didaktischen Dimension.

Die Auswirkung der Leistungsfähigkeit auf die Beurteilung der Lehre ist jedoch nicht besonders hoch. Nimmt man die Vordiploms- bzw. Zwischenprüfungsnote als Prädiktorvariable für die Bewertung der Lehr- und Studienqualität, so können ca. acht Prozent der Varianz der Lehrbeurteilungen aus der Vordiplomsnote vorhergesagt werden. In einzelnen Fächern könnte dieser Wert etwas höher liegen.

Können Vordiploms- und Zwischenprüfungsnoten als von den Hochschulen attestierte fachspezifische Leistungsfähigkeit betrachtet werden, so erscheint es von Interesse Abiturnoten, als Hinweis auf eher fachübergreifende, allgemeinere Leistungsfähigkeiten, mit den späteren Lehrbewertungen und Studienleistungen zu vergleichen.

Betrachtet man über alle Fächer hinweg die Zusammenhänge zwischen der *Abiturnote* und den Dimensionen der Lehrbeurteilung, so sind keine statistisch bedeutsamen Unterschiede festzustellen. Fachspezifisch sind etwas höhere Zusammenhänge zu erwarten (vgl. Tabelle 18).

Tabelle 18							
Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach der Abiturnote							
(Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)							
Dimensionen der Lehrqualität		Abiturnote					
		1,0 - 1,4	1,5 - 1,9	2,0 - 2,4	2,5 - 2,9	3,0 - 3,5	3,6 - 4,4
	Untergruppen n =	1074	1710	2273	1815	1233	262
	Untergruppen %	12,8	20,4	27,2	21,7	14,7	3,1
Interaktion	Mittelwert	50,2	47,8	47,6	46,69	46,79	46,4
	Standardabweichung	20,3	19,6	19,9	19,7	19,96	19,79
Didaktik	Mittelwert	53,9	51,9	51,5	50,27	50,1	49,94
	Standardabweichung	14,8	14,6	14,6	15,26	15,11	15,9
Struktur	Mittelwert	50,1	48,0	47,6	46,62	46,17	47,5
	Standardabweichung	13,8	14,6	14,3	14,63	14,64	13,91

Quelle: Konstanzer Projekt Studiensituation: Frage 7: „Welche Durchschnittsnote hatten Sie in dem Abschlußzeugnis, das Sie zur Aufnahme eines Studiums berechtigt?“

Eine andere Möglichkeit, leistungsstarke und leistungsschwache Studierende zu unterscheiden, ist ihre Selbsteinschätzung, inwieweit sie sich durch die Leistungsanforderungen persönlich belastet fühlen. Die zugrundeliegende Vermutung ist, daß sich leistungsstarke Studierende weniger belastet fühlen als eher leistungsschwache. Insgesamt kann angenommen werden, daß je stärker sich Studierende durch Leistungsanforderungen in ihrem Fach persönlich belastet fühlen, desto schlechter beurteilen sie die Lehre (vgl. Tabelle 19).

Tabelle 19				
Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach persönlicher Belastung durch Leistungsanforderungen im Fachstudium				
(Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)				
Dimensionen der Studienqualität		Belastung durch Leistungsanforderungen (Skala von 0 (= überhaupt nicht belastet) bis 6 (stark belastet))		
		gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)
	n=	1275	5062	2089
	Untergruppen %	15,1	60,1	24,8
Interaktion	Mittelwert	53,3	48,1	43,1
	Standardabweichung	20,9	19,2	19,8
Didaktik	Mittelwert	53,2	51,8	49,4
	Standardabweichung	16,3	14,4	15,0
Struktur	Mittelwert	47,9	47,9	46,7
	Standardabweichung	15,4	14,0	15,0

Quelle: Konstanzer Projekt Studiensituation: Frage 72.1: „Und in wieweit fühlen Sie sich persönlich belastet durch die Leistungsanforderungen im Fachstudium?“

Die besonders belasteten Studierenden beurteilen im Vergleich zu den nicht belasteten die Lehre signifikant schlechter, allerdings nur was die Betreuung und Beratung betrifft (Dimension Interaktion). Bezüglich der Dimensionen Didaktik und Struktur können jedoch keinen statistisch relevanten Unterschiede festgestellt werden. Betrachtet man die Lehrbeurteilung

der Gruppe besonders belasteter Studierender fachgruppenspezifisch, so werden durchaus Unterschiede auch in diesen beiden Dimensionen erkennbar (vgl. im Anhang, Tabelle A3).

Diese Unterschiede werden bereits an der Anzahl der Studierenden deutlich, die sich durch die Anforderungen in ihrem Studienfach besonders belastet fühlen. In der Medizin und in den Ingenieurwissenschaften an den Fachhochschulen sind dies mehr als ein Drittel, in der Rechtswissenschaft, den Wirtschaftswissenschaften (Uni), den Natur- und Ingenieurwissenschaften (FH) sind es jeweils um die 28 bis 29 Prozent. Bei den Wirtschaftswissenschaften (FH) gilt dies für ca. 22 Prozent, in den Kultur- und Sozialwissenschaften für ungefähr 12 Prozent, in den Sozialwissenschaften an den Fachhochschulen nur noch für 7 Prozent.

Aus diesen Verteilungen kann geschlossen werden, daß sich die empfundene Belastung nicht gleichmäßig auf die Beurteilung der Lehre durchschlägt. Ebenfalls kann nicht gesagt werden, daß dies für Fächer mit höheren Leistungsanforderungen gilt. Statt dessen findet sich in den Ergebnissen ein differenziertes Muster. Am überraschendsten zeigt sich die Auswirkung empfundener Überlastung in den Sozialwissenschaften (FH). Dort geben zwar nur insgesamt 17 Studierende an, daß sie sich besonders belastet fühlen. Ihre Lehrbeurteilungen unterscheiden sich auch nur an einer einzigen Stelle: Sie beurteilen die Betreuung und Beratung durch die Lehrenden schlechter, dies aber außerordentlich deutlich.

In den Wirtschaftswissenschaften (Uni) und den Naturwissenschaften wird die Lehre von besonders belasteten Studierenden in allen Bereichen schlechter bewertet. In der Rechtswissenschaft und den Ingenieurwissenschaften (Uni) sinken die Werte in den Dimensionen Betreuung und Beratung und Didaktik bedeutsam; in den Kulturwissenschaften, der Medizin, der Wirtschafts- und Ingenieurwissenschaften an Fachhochschulen nur im Bereich der Betreuung und Beratung.

Die vorliegenden Daten lassen die Vermutung zu, daß durch die Leistungsanforderungen in ihrem Fach besonders belastete Studierende die Studien- und Lehrqualität nicht insgesamt anders wahrnehmen, sondern vermutlich nur in einzelnen Bereichen (z.B. Beratung und Betreuung).

Als besonders belastet können jene Studierenden gelten, die sich ernsthaft Sorgen machen, das Studium überhaupt zu schaffen. Dies sind in der Stichprobe des Studierendensurveys immerhin 18 Prozent.

Tabelle 20 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach Sorgen bezüglich der Bewältigung des Studiums (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)				
Dimensionen der Studienqualität		Sorgen, das Studium überhaupt zu schaffen (Skala von 0 = trifft überhaupt nicht zu bis 6 = trifft voll und ganz zu)		
		nicht (0 - 1)	teilweise (2 - 4)	völlig (5 - 6)
	n =	3644	3260	1518
	Untergruppen %	43,3	38,7	18,0
Interaktion	Mittelwert	50,4	46,9	42,4
	Standardabweichung	20,2	18,9	19,9
Didaktik	Mittelwert	53,0	51,1	48,0
	Standardabweichung	15,0	14,4	15,0
Struktur	Mittelwert	49,8	46,7	43,9
	Standardabweichung	14,0	14,1	15,5

Quelle: Konstanzer Projekt Studiensituation: Frage 25.1: „In wie weit treffen folgende Aussagen über Lernen und Studieren auf Sie persönlich zu: Ich mache mir oft Sorgen, ob ich mein Studium überhaupt schaffe?“

Studierende, die sich durch das Studium oder andere Faktoren so belastet fühlen, daß sie sich Sorgen machen, es überhaupt zu schaffen, beurteilen die Lehre in allen Bereichen schlechter als ihre Kommilitonen, die derartige Sorgen nicht haben (vgl. Tabelle 20).

Fachgruppenspezifisch gesehen ergeben sich einige überraschende Konstellationen (siehe Anhang 2, Tabelle A5). Im allgemeinen liegt die Anzahl der Studierenden, die sich Sorgen machen, das Studium überhaupt zu schaffen, unterhalb der Anzahl besonders belasteter Studierender. Dabei sind jedoch drei Ausnahmen zu finden: In den Kulturwissenschaften und in den Sozialwissenschaften an Universitäten wie an Fachhochschulen (also in Fächern mit eher geringerer Leistungsanforderung) liegt der Anteil dieser Studierenden höher.

Betrachtet man die Auswirkungen der Sorgen hinsichtlich der Studienbewältigung auf die Lehrqualität, so ergibt sich wieder ein interessantes Muster. In den Kulturwissenschaften, den Ingenieurwissenschaften (Uni), den Wirtschaftswissenschaften und der Rechtswissenschaft beurteilen die Studierenden, die sich Sorgen um ihren Studienabschluß machen, alle Dimensionen der Lehrqualität bedeutsam schlechter als diejenigen, die sich keine Sorgen machen. In den Ingenieurwissenschaften (FH) wirken sich diese Sorgen auf die Beurteilung der Betreuung und Beratung und der Didaktik aus, in dem Fachbereich Medizin auf die Betreuung und die Bewertung der Struktur des Studiengangs. In den Sozialwissenschaften (Uni und FH) und den Naturwissenschaften wird die Dimension Interaktion kritischer gesehen. Für die Wirtschaftswissenschaften (FH) sind keine Unterschiede feststellbar.

Wurde bisher das Augenmerk auf Studierende gerichtet, die ihr Studium als mehr oder weniger problematisch empfanden, so ist festzustellen, daß eine empfundene Unterforderung sich noch deutlicher auf die Beurteilung mancher Lehraspekte auswirkt (vgl. Tabelle 21).

Tabelle 21 Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach subjektiv empfundener Charakteristik des Studiums durch hohe Leistungsnormen und -ansprüche (Mittelwerte und Standardabweichung auf einer Skala von 0 = sehr schlecht bis 100 = sehr gut)								
Dimensionen der Studienqualität		Studium charakterisiert durch hohe Leistungsnormen						
		überhaupt nicht						sehr stark
n =	Untergruppen %	202	367	750	1706	2005	2120	1261
		2,4	4,4	8,9	20,3	23,8	25,2	15,0
Interaktion	Mittelwert	47,4	47,5	46,7	49,8	50,3	46,8	42,5
	Standardabweichung	21,4	19,7	18,8	18,9	19,2	19,8	21,5
Didaktik	Mittelwert	41,5	47,2	47,4	51,6	53,8	52,5	50,7
	Standardabweichung	16,4	15,6	13,8	14,5	14,3	14,5	16,1
Struktur	Mittelwert	37,9	39,0	42,8	46,4	49,3	49,8	49,3
	Standardabweichung	14,5	14,6	13,7	13,5	13,4	14,0	16,0

Quelle: Konstanzer Projekt Studiensituation: Frage 18.2: „Wie stark ist Ihr Hauptstudienfach an Ihrer Hochschule aus Ihrer Sicht charakterisiert durch hohe Leistungsnormen, -ansprüche?“

Eine kleine Minderheit von Studierenden hat überhaupt nicht das Gefühl, daß von ihnen viel Leistung erwartet wird. Diese Studierende beurteilen den Aufbau ihres Studiengangs und die Lehrveranstaltungen signifikant schlechter als ihre Mitstudenten- und Studentinnen. Anzunehmen ist, daß sie sich schlicht unterfordert fühlen.

Am stärksten ist dies der Fall in den Sozialwissenschaften: an den Fachhochschulen für zehn Prozent, an den Universitäten für sieben Prozent der Studierenden dieser Fachgruppen. In den Kulturwissenschaften gilt dies immerhin noch für fünf Prozent der Eingeschriebenen. In den anderen Fächergruppen geht kaum ein Student, eine Studentin davon aus, daß der Studiengang überhaupt nicht durch hohe Leistungsnormen und -ansprüche zu charakterisieren wäre (die Werte schwanken zwischen einem halben und anderthalb Prozent).

Auch Studierende, die in sehr hohem Maße den Eindruck haben, in ihre Studiengang herrschen hohe Leistungsnormen -und ansprüche, beurteilen die Interaktion mit den Lehrenden tendenziell schlechter als andere Studierende.

Schlußfolgerung

Es konnte nicht eindeutig ermittelt werden, ob schlechter beurteilte Studierende tatsächlich ihrerseits schlechtere Bewertungen zur Studien- und Lehrqualität abgeben. Allerdings unterstützen die Ergebnisse tendenziell diese These.

Auch konnten einige Hinweise dafür gefunden werden, daß leistungsstärkere Studierende die Lehre ihrerseits besser einschätzen. Dies konnte an Zwischenprüfungs- bzw. Vordiplomsnoten belegt werden. Die Zusammenhänge sind jedoch nicht allzu groß, in lern- und arbeitsintensiven Studiengängen aber etwas deutlicher.

Studierende, die sich durch die Leistungsanforderungen in ihrem Fach besonders belastet fühlen, beurteilen die Lehre in der Dimension Betreuung und Beratung erheblich schlechter als die weniger belasteten Studierenden. Dies gilt insbesondere für diejenigen, die sich Sorgen machen, das Studium überhaupt bewältigen zu können.

Insgesamt kann, auch unter Einbezug fachgruppenspezifischer Besonderheiten, gesagt werden, daß leistungsschwächere Studierende die Lehre schlechter beurteilen. Dies gilt insbesondere für die Betreuung und Beratung durch die Lehrenden.

Dies heißt jedoch nicht, daß es sich hierbei um eine Verzerrung aufgrund eines unzureichenden Instrumentes handelt. Anzunehmen ist vielmehr, daß leistungsstärkere Studierende besser mit ihrem Studium zurechtkommen, dargebotene Inhalte eher verstehen und den Veranstaltungen insgesamt besser folgen können. Sie lernen vermutlich in kürzerer Zeit mehr und sind daher insgesamt mit der Lehre zufriedener. Ihnen 'bringt' die Lehre sozusagen mehr und sie schätzen sie daher auch besser ein.

Damit die unterschiedliche Leistungsstärke keine verzerrende Auswirkung auf die studentische Veranstaltungskritik ausüben kann, müßte sie theoretisch bei Befragungen miterhoben und als zu kontrollierender Einflußfaktor berücksichtigt werden. Motivation und Leistung sind jedoch keine unabhängigen Konstrukte, sondern beeinflussen sich gegenseitig. Inwieweit es tatsächlich notwendig ist, sowohl Fach- bzw. Kursinteresse als auch Leistungsfähigkeit zu kontrollieren ist von der jeweiligen Fragestellung und dem Untersuchungsziel abhängig.

10 Überfüllung von Lehrveranstaltungen

Ein Sachverhalt, der in den USA zwar häufig untersucht wurde, dort aber bei weitem nicht so relevant erscheint wie in Deutschland, ist die Kursgröße. Ein Problem amerikanischer Untersuchungen ist die Übertragbarkeit der Ergebnisse, denn Massenkurse wie sie in einigen Fächern in Deutschland üblich sind, sind in den Staaten nahezu unbekannt. Die Definition 'Großer Kurs' fängt dort schon bei der Mitgliederzahl einer großen Schulklasse an. Die Situation an nordamerikanischen Hochschulen erscheint in dieser Hinsicht wenig mit der deutschen Situation vergleichbar, weshalb zum Einfluß der Kursgröße keine Untersuchungsergebnisse aufgeführt werden sollen.

Aufgrund der Daten des Studierendensurvey können wir vor allem einen Zusammenhang zwischen der erfahrenen Überfüllung von Lehrveranstaltungen und der Lehrbewertung feststellen (vgl. Tabelle 22).

Tabelle 22

Korrelationstabelle: Überfüllung der Hochschule und Qualität von Interaktion, Didaktik und Struktur

Variable	Skalierung	Mittelwert	Standardabweichung	Besetzungszahl (N)
1. Überfüllung	0 (= überhaupt nicht) bis 6 (=sehr stark)	2,0	2,2	8421
2. Interaktion	1 (= sehr schlecht) bis 100 (= sehr gut)	47,6	19,9	8455
3. Didaktik	1 (= sehr schlecht) bis 100 (= sehr gut)	51,4	14,9	8444
4. Struktur	1 (= sehr schlecht) bis 100 (= sehr gut)	47,6	14,5	8457

N = 8457	1. Überfüllung	2. Interaktion	3. Didaktik	4. Struktur
1. Überfüllung	1,00			
2. Interaktion	- 0,32	1,00		
3. Didaktik	- 0,26	0,62	1,00	
4. Struktur	- 0,22	0,34	0,46	1,00

Quelle: Konstanzer Projekt Studiensituation: Frage 18.11: „Wie stark ist Ihr Hauptstudienfach an Ihrer Hochschule aus Ihrer Sicht charakterisiert durch Überfüllung der Lehrveranstaltungen?“

Insgesamt wirkt sich die Überfüllung von Lehrveranstaltungen auf die Beurteilung aller Dimensionen der Lehre aus, und zwar in einem Umfang, der beachtlich ist. Wichtig erscheint daher, bei dem Vergleich von Lehrveranstaltungen bzw. der Evaluation von Fachbereichen den Faktor „Studentenzahlen“ und „Überfüllung“ genügend zu berücksichtigen.

In den Fachmonographien zum Studium der Medizin wie zum Studium der Rechtswissenschaft aus studentischer Sicht, die auf der Grundlage des Studierenden surveys erarbeitet wurden, ist daher das erfahrene Ausmaß der „Überfüllung“ erfaßt und in seinen Folgen für die Beurteilung der Lehr- und Studienqualität ausgewiesen worden (vgl. Bargel/Ramm 1995; Bargel/Multrus/Ramm 1996).

11 Zusammenfassung und Folgerungen

Die Beurteilung und Bewertung der Lehre durch Studierende ist immer noch in der Diskussion. Werden einerseits mehr und mehr solcher Veranstaltungskritiken und Messungen zur Studienqualität durchgeführt, gilt es andererseits immer noch als ungeklärt, ob Studierende überhaupt willens und in der Lage sind, entsprechende Urteile abzugeben. Vermutet wird, daß Studierende dies weder sachgerecht noch exakt genug können, daß verschiedene Einflußfaktoren verzerrend auf die Lehrbeurteilung einwirken und daß die Bewertung der Studierenden alles in allem ungerecht ausfällt.

Im vorliegenden Bericht wurde die Beurteilung der Lehr- und Studienqualität von über 8.000 Studierenden hinsichtlich möglicher Störvariablen - Einflußfaktoren oder Verzerrungen - untersucht. Punkt für Punkt wurden Argumente gegen die studentische Veranstaltungskritik auf-

gegriffen und anhand der Daten des Studierendensurveys der Arbeitsgruppe Hochschulforschung geprüft.

Vielleicht am häufigsten wird das Argument vorgebracht, daß Studierende nicht über die notwendige Sachkenntnis verfügen, um die Lehrqualität richtig einschätzen zu können. Wir haben überprüft, ob Studierende mit einer höheren Sachkenntnis (erworben durch eine Ausbildung, die fachlich mit dem Studium zusammenhängt, einen Hochschulwechsel, ein Auslandsstudium, die Mitarbeit an einem Forschungsprojekt, eine Tutoriats- oder Hilfskraftstelle) sich in irgendeiner Weise in ihrer Lehrbewertung von anderen Studierenden unterscheiden. Dies ist nicht der Fall. Auch mit einer längeren Verweildauer an der Hochschule bzw. dem Abschluß des Grundstudiums ändert sich die Beurteilung der Lehr- und Studienqualität nicht.

Einem weiteren, fast ebenso häufig genannten Argument zufolge, seien Studierende zu unerfahren und unreif, um die Lehre überhaupt beurteilen zu können. Dieser Vorwurf hängt zum einen eng mit der vermuteten fehlenden Sachkenntnis zusammen, zum anderen werden Studierende als nicht erwachsen und daher als nicht 'ernsthaft' genug betrachtet. Es ist relativ schwierig, ein Konstrukt wie 'Erwachsensein' bzw. 'Reife' zu operationalisieren; aufgrund der Datenlage wurde ein Indikator, nämlich das Alter, ausgewählt. Die vorliegenden Untersuchungsergebnisse ergeben keine Differenzen in der Lehrbewertung von jüngeren und älteren Studierenden des gleichen Semesters.

Auch wurden keine geschlechtsspezifischen Unterschiede im Beurteilen von Studium und Lehrveranstaltungen ermittelt. Nur in einigen ingenieurwissenschaftlichen Studiengängen bewerten die Studentinnen die Lehrqualität schlechter als ihre Kommilitonen. Es ist allerdings aufgrund der geringen Zahl an Fächern und der ansonsten übereinstimmenden Beurteilung davon auszugehen, daß Studentinnen dort tatsächlich auf andere Situationen stoßen als ihre Mitstudenten.

Soweit zu jenen Argumenten, die die Gültigkeit und Genauigkeit studentischer Veranstaltungskritik bezweifeln, für die aber in den Angaben und Urteilen der Befragten des Studierendensurveys keine Belege zu finden waren. Anders sieht dies für die Bereiche unterschiedlicher Interessen oder unterschiedlicher Leistungsbereitschaft und Leistungsfähigkeit aus.

Die vorgenommenen Analysen zu diesen Bereichen erlauben die Vermutung, daß je leistungsmotivierter ein Student, eine Studentin ist, je stärker das Interesse am Fach bzw. am Kurs ist und je bessere die Leistungen sind, desto besser wird die Lehre beurteilt. Statistisch bedeutsame Unterschiede finden sich jedoch immer nur zwischen den Extremgruppen der sehr interessierten, motivierten oder leistungsfähigen Studierenden im Vergleich zu den wenig interessierten, motivierten und leistungsfähigen. Es ist nicht anzunehmen, daß diese drei Faktoren trennscharf voneinander sind, vielmehr ist zwischen ihnen eine hohe Abhängigkeit

zu vermuten. Leistungsmotiviertere Menschen suchen eher herausfordernde Situationen auf und bestehen diese besser. Das Interesse für ein Fach oder eine Tätigkeit hängt zudem mit dem vorherigen Können zusammen: Je höher das ist, um so höher ist meist das Interesse.

Aus diesen Gründen erscheint es nicht notwendig, alle drei möglichen Einflußfaktoren zu erheben, um sie später kontrollieren zu können. Ob es allerdings ausreichend ist, nur das entsprechende Interesse für eine Veranstaltung oder nur die fachspezifische Leistungsfähigkeit zu erheben, kann aufgrund der vorliegenden Daten nicht entschieden werden. Es ist allerdings nicht davon auszugehen, daß es sich hierbei um ein verzerrendes Bias handelt. Anzunehmen ist vielmehr, daß interessiertere und leistungsfähigere Studierende besser mit ihrem Studium zurechtkommen, mehr dafür arbeiten und dadurch besser den Veranstaltungen folgen können. Da sie die Inhalte besser rezipieren und verstehen, beurteilen sie ihrerseits folglich ganz valide die Lehrenden besser.

Was ebenfalls die Beurteilung von Lehrveranstaltung beeinträchtigen dürfte, ist der Faktor „Überfüllung“. Anhand der vorliegenden Daten des Studierendensurveys ließ sich ermitteln, daß Fachbereiche, in denen die Lehrveranstaltungen häufig überfüllt sind, insgesamt schlechtere Beurteilungen in der Studien- und Lehrqualität erzielen. Anzunehmen ist daher, daß sich der Faktor ebenfalls auf die Beurteilung der Lehrenden in ‘Massenveranstaltungen’ niederschlägt. Die Bewertungen von Lehrenden in Veranstaltungen, die sich hinsichtlich Studierendenzahl und „Überfüllung“ sehr unterscheiden, sind daher nur bedingt vergleichbar.

Aufgrund der vorgestellten Daten kann die Frage, ob studentische Veranstaltungskritiken zutreffend sind, bejaht werden. Allerdings sind zwei Punkte zu bedenken: Zum einen sind Studierende unterschiedlich interessiert und leistungsfähig. Zum anderen sind Fachbereiche und Veranstaltungen aufgrund Größe und Zusammensetzung der Studierenden nur bedingt vergleichbar.

Will man angemessene Vergleiche vornehmen, so sind die Voraussetzungen der Studierenden ebenso wie die Ausgangslage in den Fächern zu berücksichtigen. Im Konkreten heißt dies, stets zu überprüfen, ob die Beurteilung der Lehre und der Studienqualität variiert mit:

- dem Geschlecht und der Semesterzahl der Studierenden;
- dem Interesse, dem Leistungsstand und der Leistungsorientierung;
- der Studierendenzahl und Überfüllung im Fachbereich.

Allerdings ist es zudem nötig, Analysen möglichst differenziert vorzunehmen, sei es nach einzelnen Fächern, sei es nach den verschiedenen Dimensionen und Elementen der Studienqualität oder Lehrsituation.

Literatur

- Aleamoni, L.M. und Hexner, Pamela (1980). A Review of the Research on Student Evaluation and a Report on the Effect of Different Sets of Instructions on Student Course and Instructor Evaluation.. *Instructional Science*, 9, S. 67-84
- Ashton, R. H. (1975). Correlates of Ratings of Teaching Effectiveness Gordon's Survey of Interpersonal Values. *Psychological Reports*, 36, S.890
- Bargel, T. (1995). Studentische Erwartungen an die Lehr- und Studienqualität - Thesen und Erläuterungen. In: Schmitz, W. (Hg.): Evaluation der Lehre - ein Kolloquium an der Technischen Universität Dresden. Dresden, 97-104
- Bargel, T. & Ramm, M. (1994). *Das Studium der Medizin. Eine Fachmonographie aus studentischer Sicht*. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (Hrsg.), Bonn
- Bargel, T., Multrus, F. & Ramm, M. (1996). *Das Studium der Rechtswissenschaft. Eine Fachmonographie aus studentischer Sicht*. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (Hrsg.), Bonn
- Bargel, T., Multrus, F. & Ramm, M. (1996). *Studium und Studierende in den 90er Jahren. Entwicklung an Universitäten und Fachhochschulen in den alten und neuen Bundesländern*. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (Hrsg.), Bonn
- Blass, T. (1980). What Do Positive Correlations Between Student Grades and Teacher Evaluation Mean? *Teaching of Psychology*, 7 (3), S. 186-187
- Bledsoe, J. C. (1978). Insight into one's Own Teaching: Stability of Students' Evaluations Across Classes. *Psychological Reports*, 42, S. 1071-1074
- Blount, H. P., Stallings, W. M. und Gupta, V. G. (1978). The Effects of Different Instructions on Student Ratings of University Courses and Teachers. *Journal of Educational Research*, 71, S.149-152
- Brandenburg, D. C., Slinde, J. A. und Batista, E. E. (1977). Student Ratings of Instruction: Validity and Normative Interpretation. *Research in Higher Education*, 7, S. 67-78
- Brown, D. L. (1976). Faculty Ratings and Student Grades: A University-wide Multiple Regression Analysis. *Journal of Educational Psychology*, 68 (5), S. 573-576
- Bruton, B. T. und Crull, Sue R. (1982). Causes and Consequences of Student Evaluation of Instruction. *Research in Higher Education*, 17 (3), S. 195-206
- Centra, J. A. (1977). Student Ratings of Instruction and Their Relationship to Student Learning. *American Educational Research Journal*, 14, S. 17-24
- Cooper, Pamela J.; Steward, Lea P. und Gudykunst, W. B. (1982). Relationship with Instructor and other Variables Influencing Student Evaluations of Instruction. *Communication Quarterly*, 30 (4), S.308-314
- Daniel, H.-D. (1994). Hörerbefragung an der Universität Mannheim: Konzeption, Erhebung, Auswertung. *Empirische Pädagogik*, 1994, 8 (2), S. 109-130

- Daniel, H.-D. (1995). Das Modellprojekt „Evaluation der Lehre“ an der Universität Mannheim. Teil 2: Statistische Auswertung von Befragungen in Lehrveranstaltungen. In Mohler, P. *Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung*. Münster: Warmann
- Diehl, J.M. (1994). *Fragebogen zur studentischen Evaluation von Hochschulveranstaltungen. Manual zum VBVOR und VBREF*. Fachbereich 06 Psychologie (Abteilung Methodik), Universität Gießen.
- Driscoll, Laura A. und Goodwin, W. L. (1979). The Effects of Varying Information About Use and Disposition of Results on University Students' Evaluation of Faculty and Courses. *American Educational Research Journal*, 16, S. 25-37
- DuCette, J. und Kenney, Jane (1982). Do Grading Standards Affect Student Evaluations of Teaching? Some New Evidence on an Old Question. *Journal of Educational Psychology*, 3, S. 308-314
- Elmore, Patricia B. und Pohlmann, J. T. (1978). Effect of Teacher, Student, and Class Characteristics on the Evaluation of College Instructors. *Journal of Educational Psychology*, 70, S.187-192
- Gaffuri, Adriane, Wrench, D., Karr, C. und Kopp, R. (1982). Exploring Some Pitfalls in Student Evaluation of Teaching. *Teaching of Psychology*, 9 (4), S.229-230
- Gawatz, R. (1991). *Studium - Wissenschaft - Beruf. Berufliche Studienperspektive westdeutscher Studierender und ihr Stellenwert für die Studienbewältigung und Studiensituation*. Konstanz: Hartung-Gorre-Verlag
- Hage, N. el (1995). *Zur Validität studentischer Veranstaltungskritiken. Befunde empirischer Studien zu einem Umstrittenen Verfahren*. Hefte zur Bildungs- und Hochschulforschung, Bd. 13
- Hage, N. el (1996). *Lehrevaluation und studentische Veranstaltungskritik*. Projekte, Instrumente und Grundlagen. Bonn: Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie.
- Hofman, J. M. (1988). Studienmotivation und Veranstaltungsbeurteilung. *Psychologie, Erziehung, Unterricht*, 35, S.119-126
- Hornbostel, S. und Daniel, H.-D. (1993) Das Spiegel-Ranking: Mediensensation oder ein Beitrag zur hochschulvergleichenden Lehrevaluation? In: Gralki, H.D., Grünh, D. & Hecht, Heidemarie. *Evaluation schafft Autonomie. Perspektiven der Lehrbewertung an Hochschulen*
- Kovac, Robert (1976). Personality Correlates of Faculty and Course Evaluations. *Research in Higher Education*, 5, S. 335-344
- Kromrey, H. (1993). Studentische Vorlesungskritik. Empirische Daten und Konsequenzen für die Lehre, *Soziologie*, 1, 39-56
- Kromrey, H. (1994). Wie erkennt man gute Lehre? Was studentische Vorlesungsbefragungen (nicht) aussagen. *Empirische Pädagogik*, 2, S. 153-168
- Kromrey, H. (1995). Evaluation der Lehre durch Umfrageforschung? Methodische Fallstricke bei der Messung von Lehrqualität durch Befragung von Vorlesungsteilnehmern. In: Moh-

- ler, P. *Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung*. Münster: Waxmann
- Leitow, Bettina (1996). Studentische Haltungen zur Studieneffizienz - Ein Beitrag zur Typologie studentischer Orientierungen, Hefte zur Bildungs- und Hochschulforschung, Bd. 18
- Lester, D. (1982). Students' Evaluation of Teaching and Course Performance. *Psychological Reports*, 50, S.1126
- Lind, G. (1993). *Moral und Bildung*. Heidelberg: Asanger Verlag
- Marsh, H. W. (1982). The Use of Path Analysis to Estimate Teacher and Course Effects in Student Ratings of Instructional Effectiveness. *Applied Psychological Measurement*, 6 (1), S.47-59
- Marsh, H. W. (1982b). Validity of Students' Evaluation of College Teaching. A Multitrait-Multimethod Analysis. *Journal of Educational Psychology*, 74 (2), S. 264-279
- McKeachie, W.J. und Lin, Y.G. (1978). A Note on Validity of Student Ratings of Teaching. *Educational Research Quarterly*, 4 (3), S. 45-47
- Multrus, F. (1995). *Zur Lehr- und Studienqualität. Dimensionen, Skalen und Befunde des Studierendensurveys*. Hefte zur Bildungs- und Hochschulforschung, Bd. 12,
- Orpen, C. (1980). The Susceptibility of Student Evaluation of Lecturers to Situation Variabls. *Higher Education*, 9, S. 293-306
- Peisert, H., Bargel, T. & Framheim, Gertrud (1988). *Studiensituation und studentische Orientierung an Universitäten und Fachhochschulen*. Zweite Erhebung zur Studiensituation im WS 1984/85. Bundesministerium für Bildung und Wissenschaft (Hrsg.) Schriftreihe: Studien zur Bildung und Wissenschaft, Bd. 59, Bad Honnef: Bock-Verlag
- Petchers, Marcia K. und Chow, J. C. (1988). Sources of Variation in Students' Evaluation of Instruction in a Graduate Social Work Program. *Journal of Social Work Education*, 1, S. 35-42
- Rindermann, H. (1996). *Untersuchungen zur Brauchbarkeit studentischer Lehrevaluationen*. Landau: Verlag Empirischer Pädagogik
- Rindermann, H. & Amelang, M. (1994a). *Das Heidelberger Inventar zur Lehrveranstaltungs-evaluation (HILVE)*. Heidelberg: Asanger.
- Romeo, Felicia F. und Weber, W. A. (1985). An Examination of Variables which Influence Student Ratings of University Faculty. *College Student Journal*, 19, S. 133-140
- Snyder, C.R. und Clair, M. (1976). Effects of Expected and Obtained Grades on Teacher Evaluation and Attribution of Performance. *Journal of Educational Psychology*, 68 (1), S. 75-82
- SPIEGEL-SPEZIAL (1993). *Welche Uni ist die beste?* Bd. 3, Hamburg
- Yates, J.W. und Karmos, J. S. (1971). Final Grade Predictions and Instructional Evaluation at Different Levels of Course Achievement. *Journal of Instructional Psychology*, 4(3), S. 38-40

Anhang

Fachspezifische Befunde zu Einflüssen auf die studentischen Urteile zur Lehr- und Studienqualität

Tabelle A1a und A1b Studienfächer und Ablegen einer Zwischenprüfung.....	48-49
Tabelle A2 Studienfächer und Alter sowie Studienjahr.....	50
Tabelle A3 Studienfächer und subjektive Belastung durch Leistungsanforderungen	51
Tabelle A4 Studienfächer und Leistungsstand (Note in Zwischenprüfung bzw. Vordiplom).....	52
Tabelle A5 Studienfächer und Sorgen über Bewältigung des Studiums	53

Tabelle A1a													
Beurteilung von Interaktion, Didaktik und Struktur im Studiengang in Abhängigkeit vom Studienfach (Auswahl) und dem Ablegen einer Zwischenprüfung.													
(Mittelwerte und Standardabweichung auf einer Skala von 0 (=sehr schlecht) bis 100 (=sehr gut))											gesamt N = 4091		
Dimensionen der Lehrqualität	n (Fach)=	Germanistik		Jura		Psychologie		Erz.wiss./ (Sonder-) Pädagogik		Betriebswirtschaft		Humanmedizin	
		196		400		144		264		602		488	
		nein	ja	nein	ja	nein	ja	nein	ja	nein	ja	nein	ja
Zwischenprüfung abgelegt	n =	74	122	167	233	54	90	129	135	243	359	145	343
Untergruppen %		37,8	62,2	41,8	58,3	37,5	62,5	48,9	51,1	40,4	59,6	29,7	70,3
Interaktion	Mittelwert	54,1	52,6	43,7	40,1	51,9	51,2	52,0	53,0	41,9	39,9	39,2	33,3
	Standardabweichung	19,2	18,5	20,5	20,0	18,4	17,9	16,4	19,1	18,5	18,7	18,2	16,9
Didaktik	Mittelwert	54,9	53,5	54,4	46,8	52,7	48,0	52,7	49,0	51,0	48,0	53,1	48,3
	Standardabweichung	16,2	14,5	15,3	15,5	14,1	16,2	11,9	14,4	13,5	14,3	13,7	13,8
Struktur	Mittelwert	45,2	41,6	49,3	43,4	48,0	47,7	43,0	43,5	51,4	46,0	49,7	50,3
	Standardabweichung	16,5	14,1	13,5	15,0	14,3	13,7	14,8	14,4	13,1	14,4	14,6	14,6
Dimensionen der Lehrqualität	n (Fach)=	Mathematik		Physik		Chemie		Maschinenbau		Elektro- / Nach-richtentechnik		Bauingenieurwesen	
		153		212		190		643		442		357	
		nein	ja	nein	ja	nein	ja	nein	ja	nein	ja	nein	ja
Zwischenprüfung abgelegt	n =	59	94	45	167	55	135	164	479	101	341	164	193
Untergruppen %		38,6	61,4	21,2	78,8	28,9	71,1	25,5	74,5	22,9	77,1	45,9	54,1
Interaktion	Mittelwert	50,7	56,7	51,1	50,1	48,0	49,2	46,7	50,0	49,7	50,8	44,8	45,9
	Standardabweichung	17,4	18,5	15,8	19,0	22,2	20,9	18,5	19,9	18,5	18,5	18,3	18,3
Didaktik	Mittelwert	50,8	56,0	50,3	50,6	53,0	51,2	52,2	53,9	52,0	53,7	52,0	51,5
	Standardabweichung	13,4	13,5	13,8	16,7	17,8	14,4	13,9	15,4	14,1	14,5	14,5	14,3
Struktur	Mittelwert	46,2	49,6	51,0	48,3	51,5	46,9	50,7	51,4	49,9	51,1	48,9	49,4
	Standardabweichung	13,6	13,8	11,3	14,8	15,0	14,9	13,2	13,7	11,5	13,4	13,1	13,3

Studierende, die ihre Zwischenprüfung nicht bestanden haben wurden aufgrund ihrer geringen Zahl nicht berücksichtigt.

Quelle: Konstanzer Projekt Studiensituation: Frage 34: „Haben Sie bereits die Zwischenprüfung/das Vordiplom in Ihrem Hauptfach abgelegt?“

Tabelle A1b													
Beurteilung von Interaktion, Didaktik und Struktur im Studiengang in Abhängigkeit vom Studienfach (Auswahl) und dem Ablegen einer Zwischenprüfung.													
(Mittelwerte und Standardabweichung auf einer Skala von 0 (=sehr schlecht) bis 100 (=sehr gut))											gesamt N = 2140		
Dimensionen der Studienqualität		Geschichte		Anglistik		Wirtschaftswissenschaft		Volkswirtschaft		Informatik		Biologie	
n (Fach)=		111		111		167		123		230		192	
Zwischenprüfung abgelegt		nein	ja	nein	ja	nein	ja	nein	ja	nein	ja	nein	ja
n =		37	74	42	69	70	97	47	76	86	144	61	131
Untergruppen %		33,3	66,7	37,8	62,2	41,9	58,1	38,2	61,8	37,4	62,6	31,8	68,2
Interaktion	Mittelwert	61,3	57,2	52,2	55,4	39,1	34,0	40,1	37,5	47,2	45,4	48,2	45,8
	Standardabweichung	18,6	18,3	19,7	19,7	18,3	18,5	17,9	18,6	16,5	19,1	15,8	19,6
Didaktik	Mittelwert	59,0	55,0	53,9	54,5	47,1	45,6	48,9	45,7	49,5	49,7	52,3	51,1
	Standardabweichung	15,5	15,1	12,2	14,7	16,9	13,2	14,7	14,6	14,0	14,8	12,1	14,8
Struktur	Mittelwert	44,2	44,1	47,4	43,2	45,9	41,2	47,1	42,5	49,6	48,2	50,8	46,0
	Standardabweichung	13,0	13,0	14,9	13,5	15,4	15,0	14,0	14,0	14,9	14,1	12,5	13,8
Studienfach (nur Universitäten)		Betriebswirtschaft		Maschinenbau		Elektro- / Nachrichtentechnik		Architektur		Bauingenieurwesen		Studierende gesamt	
n (Fach)=		382		320		200		104		191		7568	
Zwischenprüfung abgelegt		nein	ja	nein	ja	nein	ja	nein	ja	nein	ja	nein	ja
n =		141	241	65	255	46	154	38	66	87	104	2775	4793
Untergruppen %		36,9	63,1	20,3	79,7	23,0	77,0	36,5	63,5	45,5	54,5	36,7	63,3
Interaktion	Mittelwert	38,3	36,8	43,5	47,7	46,3	48,3	50,6	45,0	40,5	41,9	48,8	47,2
	Standardabweichung	17,4	17,6	18,5	20,6	19,5	19,6	17,7	18,8	17,0	16,3	19,4	20,1
Didaktik	Mittelwert	49,2	46,4	49,1	52,7	50,1	50,9	51,3	46,5	48,8	49,8	52,9	50,8
	Standardabweichung	12,9	14,3	13,7	14,9	13,5	15,2	15,7	13,1	13,6	12,8	14,4	15,1
Struktur	Mittelwert	50,1	44,2	48,1	52,3	51,9	52,8	47,8	46,3	47,9	49,7	48,5	47,5
	Standardabweichung	12,8	14,8	12,4	13,9	12,2	13,1	13,7	13,7	13,1	12,5	14,2	14,5

Studierende, die ihre Zwischenprüfung nicht bestanden haben wurden aufgrund ihrer geringen Zahl nicht berücksichtigt.

Quelle: Konstanzer Projekt Studiensituation: Frage 34: „Haben Sie bereits die Zwischenprüfung/das Vordiplom in Ihrem Hauptfach abgelegt?“

Tabelle A2													
Beurteilung von Interaktion, Didaktik und Struktur im Studiengang fachspezifisch in Abhängigkeit des Alters und Studienjahres der Befragten.													
(Mittelwerte und Standardabweichung auf einer Skala von 0 (=sehr schlecht) bis 100 (=sehr gut))													
													gesamt N = 8364
Dimensionen der Studienqualität	n (Fach) =	1. Studienjahr				2. Studienjahr				3. Studienjahr			
	% von ges.N =	933 11,2				1076 12,9				1336 16,0			
Alter		bis 20	21-25	26-30	über 30	bis 20	21-25	26-30	über 30	bis 20	21-25	26-30	über 30
	n =	418	428	66	21	223	715	101	37	49	1034	204	49
	Untergruppen %	44,8	45,9	7,1	2,3	20,7	66,4	9,4	3,4	3,7	77,4	15,3	3,7
Interaktion	Mittelwert	52,9	51,7	55,0	60,1	53,1	47,6	52,1	57,1	51,7	47,1	47,9	49,3
	Standardabweichung	17,4	18,7	20,2	20,1	20,4	19,4	19,2	20,5	17,5	19,7	19,5	19,5
Didaktik	Mittelwert	56,5	55,5	56,1	56,1	55,8	52,6	52,2	54,0	55,0	52,2	51,4	51,8
	Standardabweichung	13,2	14,5	15,7	17,3	14,3	13,5	14,9	15,5	12,4	14,7	14,2	14,2
Struktur	Mittelwert	52,1	49,6	49,9	50,5	50,9	49,3	49,9	47,9	47,5	48,7	46,9	48,2
	Standardabweichung	13,2	14,3	15,2	19,3	14,8	12,9	12,8	13,4	13,1	14,3	13,8	14,6
Dimensionen der Studienqualität	n (Fach) =	4. Studienjahr				5. Studienjahr				über 5. Studienjahr			
	% von ges.N =	1506 18,0				1283 15,3				2230 26,7			
Alter		bis 20	21-25	26-30	über 30	bis 20	21-25	26-30	über 30	bis 20	21-25	26-30	über 30
	n =	4	1059	384	59	3	776	465	39	2	396	1380	452
	Untergruppen %	0,3	70,3	25,5	3,9	0,2	60,5	36,2	3,0	0,1	17,8	61,9	20,3
Interaktion	Mittelwert	63,3	48,5	46,4	48,5	45,3	48,5	44,5	45,3	48,0	45,7	44,3	45,6
	Standardabweichung	10,4	19,9	18,8	19,3	32,2	20,1	19,7	17,6	35,4	20,4	20,1	21,0
Didaktik	Mittelwert	59,0	51,4	49,4	50,6	52,7	51,8	50,3	47,6	48,0	51,3	48,1	48,8
	Standardabweichung	8,5	15,0	14,8	14,6	15,0	15,1	15,5	14,6	14,1	14,4	15,3	15,7
Struktur	Mittelwert	49,8	47,5	46,1	48,9	52,7	47,5	46,5	45,5	41,0	47,2	45,1	44,8
	Standardabweichung	25,3	14,0	14,7	15,0	5,8	15,2	14,5	12,8	7,1	14,3	15,0	14,7

Quelle: Konstanzer Projekt Studiensituation; Fragen 51 (Hochschulsemester) und 119 (Alter).

Tabelle A3

Beurteilung von Interaktion, Didaktik und Struktur im Studiengang in Abhängigkeit der subjektiv empfundenen Belastung durch die Anforderungen im Fachstudium nach Fächergruppen.

(Mittelwerte und Standardabweichung auf einer Skala von 0 (=sehr schlecht) bis 100 (=sehr gut))

gesamt N = 8022

		UNIVERSITÄTEN														
		Kulturwissenschaften 1277			Sozialwissenschaften 682			Rechtswissenschaft 633			Wirtschaftswiss. 871			Medizin 644		
Belastung (Skala von 0 (= überhaupt nicht belastet) bis 6 (stark belastet))	n (Fach)=	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)
	Dimensionen der Studienqualität															
	n =	318	795	164	176	426	80	65	389	179	85	545	241	68	338	238
	Untergruppen %	24,9	62,3	12,8	25,8	62,5	11,7	10,3	61,5	28,3	9,8	62,6	27,7	10,6	52,5	37,0
Interaktion	Mittelwert	58,2	54,9	52,2	51,7	50,4	51,4	47,2	41,7	41,0	37,4	38,3	32,2	40,0	36,9	34,4
	Standardabweichung	20,7	19,1	20,1	18,8	17,1	17,8	20,7	19,4	21,4	19,0	16,6	17,8	19,5	17,5	19,4
Didaktik	Mittelwert	54,8	54,1	52,8	48,3	50,2	51,4	56,3	50,5	48,0	48,8	48,0	43,8	51,3	50,8	50,4
	Standardabweichung	17,3	14,2	16,8	14,4	14,0	14,9	15,7	15,9	14,6	15,4	13,4	14,4	16,1	14,2	14,5
Struktur	Mittelwert	44,1	44,2	40,9	42,6	44,3	42,4	48,4	46,5	43,7	50,5	47,0	42,5	49,0	51,0	48,5
	Standardabweichung	14,5	14,1	14,9	15,5	14,2	15,3	13,8	14,3	15,7	14,7	13,5	15,8	16,4	14,3	15,2
		UNIVERSITÄTEN						FACHHOCHSCHULEN								
		Naturwissenschaften 1252			Ingenieurwissenschaften 993			Sozialwissenschaften 254			Wirtschaftswiss. 338			Ingenieurwissenschaften 1078		
Belastung (Skala von 0 (= überhaupt nicht belastet) bis 6 (stark belastet))	n (Fach)=	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)
	Dimensionen der Studienqualität															
	n =	163	726	363	123	583	287	74	163	17	44	220	74	83	609	386
	Untergruppen %	13,0	58,0	29,0	12,4	58,7	28,9	29,1	64,2	6,7	13,0	65,1	21,9	7,7	56,5	35,8
Interaktion	Mittelwert	53,8	49,3	44,1	52,5	45,9	43,4	62,5	55,3	52,2	51,2	46,1	45,3	56,1	54,3	48,7
	Standardabweichung	20,1	18,8	19,1	19,7	19,0	18,6	19,5	17,5	14,8	19,3	17,7	19,1	20,4	17,9	17,5
Didaktik	Mittelwert	54,8	51,8	48,2	53,6	51,0	48,0	54,4	51,4	53,5	54,5	51,3	51,8	56,7	55,3	52,8
	Standardabweichung	15,1	14,7	15,2	15,2	14,2	13,6	17,1	15,8	16,5	13,8	13,7	15,1	18,9	13,6	14,8
Struktur	Mittelwert	52,6	47,6	47,2	52,6	50,2	50,4	46,1	46,9	45,2	51,0	51,7	52,3	53,1	51,6	49,3
	Standardabweichung	14,5	13,6	14,6	14,8	13,3	12,9	14,6	12,5	16,3	14,5	13,1	13,3	15,1	12,6	13,8

Quelle: Konstanzer Projekt Studiensituation; Frage 72.1: „Und inwieweit fühlen Sie sich persönlich belastet durch die Leistungsanforderungen im Fachstudium?“

Tabelle A4

Beurteilung von Interaktion, Didaktik und Struktur im Studiengang in Abhängigkeit vom Studienfach und der Durchschnittsnote der Zwischenprüfung/des Vordiploms.

(Mittelwerte und Standardabweichung auf einer Skala von 0 (=sehr schlecht) bis 100 (=sehr gut))

N gesamt = 3918

Studienfach	Germanistik			Jura			Psychologie			Erz.wiss./(Sonder-) Pädagogik			Betriebswirtschaft			Humanmedizin			
	n (Fach)=	203		449			125			254			542			447			
Durchschnittsnote Zwischenprüfung	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	
n =	113	84	6	32	235	182	67	52	6	111	134	9	48	338	156	112	204	131	
Untergruppen %	55,7	41,4	3,0	7,1	52,3	40,5	53,6	41,6	4,8	43,7	52,8	3,5	8,9	62,4	28,8	25,1	45,6	29,3	
Interaktion	Mittelwert	54,7	49,8	52,0	43,0	42,9	38,4	52,3	48,9	48,7	54,4	48,9	46,9	42,3	40,6	40,7	37,2	35,0	28,3
	Standardabweichung	18,9	16,4	28,3	20,4	20,3	19,3	18,3	17,3	21,5	19,1	16,7	20,2	20,6	18,4	18,5	16,4	17,6	15,8
Didaktik	Mittelwert	54,7	51,9	54,8	53,9	49,5	47,5	50,3	47,5	44,5	49,8	48,2	49,8	50,9	48,3	49,6	52,5	49,7	44,7
	Standardabweichung	15,5	14,2	23,2	18,1	15,5	15,5	15,3	14,3	14,2	13,0	14,0	17,7	15,9	14,0	13,8	12,8	13,6	14,8
Struktur	Mittelwert	44,0	41,3	31,3	48,6	45,8	44,3	49,8	46,8	37,5	41,6	43,0	41,8	48,1	48,2	47,3	53,0	51,1	45,1
	Standardabweichung	14,5	13,7	13,3	12,8	14,9	14,6	14,0	13,2	12,2	14,2	14,3	11,1	13,4	14,1	14,9	12,7	14,5	15,6

Studienfach	Mathematik			Physik			Chemie			Maschinenbau			Elektro- / Nachrichtentechnik			Bauingenieurwesen			
	n (Fach)=	146		199			181			615			425			332			
Durchschnittsnote Zwischenprüfung	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	1,0 - 2,0	2,1 - 3,0	3,1 - 4,4	
n =	45	67	34	101	81	17	72	89	20	84	377	154	45	268	112	38	210	84	
Untergruppen %	30,8	45,9	23,3	50,8	40,7	8,5	39,8	49,2	11,0	13,7	61,3	25,0	10,6	63,1	26,4	11,4	63,3	25,3	
Interaktion	Mittelwert	58,6	54,5	46,2	51,0	47,8	54,4	51,4	46,4	45,1	52,1	49,5	46,9	52,1	51,9	45,6	46,7	45,7	42,0
	Standardabweichung	18,2	19,1	17,1	18,1	18,4	22,0	22,8	20,7	18,6	17,1	20,2	19,3	18,6	18,4	17,5	18,5	17,5	19,0
Didaktik	Mittelwert	56,2	54,2	48,8	50,3	50,0	48,3	52,8	51,1	48,4	54,8	53,5	51,4	52,8	54,2	51,0	52,7	52,1	50,0
	Standardabweichung	14,1	13,9	13,0	16,1	15,9	21,5	14,5	14,9	19,7	13,2	15,4	14,9	15,6	13,9	13,6	13,5	14,1	14,1
Struktur	Mittelwert	51,5	46,9	43,7	48,2	49,0	44,4	51,0	46,0	43,9	51,9	51,4	50,2	51,3	51,1	50,2	51,2	49,9	45,8
	Standardabweichung	12,5	15,9	11,5	15,1	13,4	11,4	15,5	14,7	13,9	13,6	13,6	13,8	13,2	13,2	12,0	12,9	12,6	12,8

Quelle: Konstanzer Projekt Studiensituation: Frage 35: „Geben Sie bitte auch die Durchschnittsnote Ihrer Zwischenprüfung/des Vordiploms an. (Wenn Sie keine Note erhalten haben oder keine Zwischenprüfung abgelegt haben, versuchen Sie bitte, Ihre bisherigen Studienleistungen als Notendurchschnitt zu schätzen.)“

Tabelle A5

Beurteilung von Interaktion, Didaktik und Struktur im Studiengang nach Sorgen bezüglich der Bewältigung des Studiums und Fachgruppen.

(Mittelwerte und Standardabweichung auf einer Skala von 0 (=sehr schlecht) bis 100 (=sehr gut))

gesamt N = 8019

Dimensionen der Studienqualität		UNIVERSITÄTEN														
		Kulturwissenschaften			Sozialwissenschaften			Rechtswissenschaft			Wirtschaftswiss.			Medizin		
n (Fach)=		1273			683			632			874			640		
Sorgen, das Studium zu schaffen (Skala von 0 (= trifft überhaupt nicht zu) bis 6 (trifft voll und ganz zu))		gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)
n =		499	515	259	298	263	122	204	276	152	359	346	169	287	241	112
Untergruppen %		39,2	40,5	20,3	43,6	38,5	17,9	32,3	43,7	24,1	41,1	39,6	19,3	44,8	37,7	17,5
Interaktion	Mittelwert	59,6	54,0	49,4	54,0	49,7	45,6	44,1	42,0	39,2	38,2	37,1	31,9	37,4	37,1	32,0
	Standardabweichung	19,7	18,6	20,2	17,9	16,8	17,1	21,0	18,8	21,0	17,5	17,1	17,2	18,0	18,2	19,6
Didaktik	Mittelwert	56,8	53,8	49,1	51,3	49,4	47,0	52,9	49,7	47,8	48,5	47,7	42,1	51,7	49,7	50,3
	Standardabweichung	15,6	14,5	15,4	14,4	14,0	14,1	17,1	14,5	14,9	13,9	13,5	14,6	14,3	14,5	14,7
Struktur	Mittelwert	46,5	43,2	38,6	45,8	42,2	41,1	48,3	45,7	42,9	49,3	45,0	41,3	52,1	49,1	46,0
	Standardabweichung	13,7	14,2	14,5	14,8	13,9	15,4	14,0	14,3	15,6	14,0	13,6	16,3	14,3	14,4	16,6

Dimensionen der Studienqualität		UNIVERSITÄTEN						FACHHOCHSCHULEN								
		Naturwissenschaften			Ingenieurwissenschaften			Sozialwissenschaften			Wirtschaftswiss.			Ingenieurwissenschaften		
n (Fach)=		1250			993			254			337			1083		
Sorgen, das Studium zu schaffen (Skala von 0 (= trifft überhaupt nicht zu) bis 6 (trifft voll und ganz zu))		gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)	gering (0 - 1)	mittel (2 - 4)	erheblich (5 - 6)
n =		564	486	200	490	347	156	142	86	26	146	135	56	450	417	216
Untergruppen %		45,1	38,9	16,0	49,3	34,9	15,7	55,9	33,9	10,2	43,3	40,1	16,6	41,6	38,5	19,9
Interaktion	Mittelwert	50,7	47,1	45,2	49,3	44,3	39,0	58,5	56,3	53,1	47,4	47,3	43,0	55,9	51,6	46,5
	Standardabweichung	19,6	18,6	19,0	18,8	18,8	18,4	19,2	16,4	18,1	18,5	17,4	19,6	18,1	17,5	17,9
Didaktik	Mittelwert	52,1	50,7	49,2	52,8	49,1	45,5	52,8	51,7	52,5	52,1	52,7	49,0	56,1	54,8	50,5
	Standardabweichung	14,9	15,2	15,0	14,1	13,9	14,1	16,9	15,9	14,4	14,1	13,7	14,9	14,7	13,8	14,8
Struktur	Mittelwert	49,7	47,1	45,7	52,6	49,1	47,0	46,7	47,1	43,9	52,2	51,4	51,0	52,6	50,5	47,9
	Standardabweichung	14,1	13,9	14,2	13,2	13,1	13,7	13,1	13,1	15,8	12,7	13,1	15,5	12,8	12,9	14,6

Quelle: Konstanzer Projekt Studiensituation: Frage 25.1: „In wie weit treffen folgende Aussagen über Lernen und Studieren auf Sie persönlich zu: Ich mache mir oft Sorgen, ob ich mein Studium überhaupt schaffe?“

ISSN 1616-0398