

Bargel, Tino; El Hage, Natalija

Evaluation der Hochschullehre. Modelle, Probleme und Perspektiven

Helmke, Andreas [Hrsg.]; Hornstein, Walter [Hrsg.]; Terhart, Ewald [Hrsg.]: *Qualität und Qualitätssicherung im Bildungsbereich; Schule, Sozialpädagogik, Hochschule. Weinheim* : Beltz 2000, S. 207-224. - (Zeitschrift für Pädagogik, Beiheft; 41)



Quellenangabe/ Reference:

Bargel, Tino; El Hage, Natalija: Evaluation der Hochschullehre. Modelle, Probleme und Perspektiven - In: Helmke, Andreas [Hrsg.]; Hornstein, Walter [Hrsg.]; Terhart, Ewald [Hrsg.]: *Qualität und Qualitätssicherung im Bildungsbereich; Schule, Sozialpädagogik, Hochschule. Weinheim* : Beltz 2000, S. 207-224 - URN: urn:nbn:de:0111-opus-84935 - DOI: 10.25656/01:8493

<https://nbn-resolving.org/urn:nbn:de:0111-opus-84935>

<https://doi.org/10.25656/01:8493>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

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

Terms of use

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

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

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Digitalisiert

Zeitschrift für Pädagogik

41. Beiheft

Zeitschrift für Pädagogik

41. Beiheft

Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule

Herausgegeben von

Andreas Helmke, Walter Hornstein und Ewald Terhart

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder ein anderes Verfahren – reproduziert oder in eine von Maschinen, insbesondere Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch Vortrag, Funk- und Fernsehsendung, im Magnettonverfahren oder auf ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen oder sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopie hergestellt werden. Jede im Bereich eines gewerblichen Unternehmens hergestellte oder genutzte Kopie dient gewerblichen Zwecken gem. § 54 (2) UrhG und verpflichtet zur Gebührenzahlung an die VG Wort, Abteilung Wissenschaft, Goethestr. 49, 80336 München, von der die einzelnen Zahlungsmodalitäten zu erfragen sind.

© 2000 Beltz Verlag · Weinheim und Basel
Herstellung: Klaus Kaltenberg
Satz: Satz- und Reprotechnik GmbH, Hemsbach
Druck: Druckhaus „Thomas Müntzer“, Bad Langensalza
Printed in Germany
ISSN 0514-2717

Bestell-Nr. 41142

Inhaltsverzeichnis

Einleitung

ANDREAS HELMKE/WALTER HORNSTEIN/EWALD TERHART

Qualität und Qualitätssicherung im Bildungsbereich.

Zur Einleitung in das Beiheft. 7

Teil I

Qualität und Qualitätssicherung – analytisch betrachtet

LEE HARVEY/DIANA GREEN

Qualität definieren. Fünf unterschiedliche Ansätze 17

HELMUT HEID

Qualität: Überlegungen zur Begründung einer pädagogischen

Beurteilungskategorie 41

Teil II

Qualität und Qualitätssicherung in der Schule

HELMUT FEND

Qualität und Qualitätssicherung im Bildungswesen:

Wohlfahrtsstaatliche Modelle und Marktmodelle 55

HARTMUT DITTON

Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht.

Ein Überblick zum Stand der empirischen Forschung. 73

HERBERT ALTRICHTER

Konfliktzonen beim Aufbau schulischer

Qualitätssicherung und Qualitätsentwicklung 93

TOM STRYCK

Qualitätssicherung in der Geisterbahn.

Was hat die Schulaufsicht mit Schulqualität zu tun? 111

Teil III

Qualität und Qualitätssicherung in der Sozialpädagogik

WALTER HORNSTEIN

Qualität und Evaluation in der Sozialpädagogik 129

MICHAEL WINKLER

Qualität und Jugendhilfe:
Über Sozialpädagogik und reflexive Modernisierung 137

JOACHIM MERCHEL

Zwischen Hoffnung auf Qualifizierung und Abwehrreflex:
Zum Umgang mit dem Qualitätsthema in der Jugendhilfe 161

RAINER WETZLER

Internationale Ansätze zur Qualitätssicherung und -verbesserung
in ausgewählten Bereichen sozialer Arbeit 185

Teil IV

Qualität und Qualitätssicherung in der Hochschule

TINO BARGEL/NATALIJA EL HAGE

Evaluation der Hochschullehre: Modelle, Probleme und Perspektiven . . . 207

WOLFF-DIETRICH WEBLER

Weiterbildung der Hochschullehrer als Mittel der Qualitätssicherung 225

BIRGIT BERENDT

Was ist gute Hochschullehre? 247

FRIEDRICH-WILHELM SCHRADER/ANDREAS HELMKE

Wirksamkeit des Hochschulunterrichts aus der Sicht der Studierenden.
Eine empirische Studie 261

Evaluation der Hochschullehre

Modelle, Probleme und Perspektiven

Die Evaluation der Hochschullehre hat sich trotz mancher Bedenken an den deutschen Universitäten und Fachhochschulen durchgesetzt, wenngleich sie noch keineswegs selbstverständlich oder alltäglich geworden ist. Die Hochschulrektorenkonferenz hat dazu eine Stellungnahme (1998), der Wissenschaftsrat (1996) Empfehlungen vorgelegt. Dennoch ist die Evaluation von Lehre und Lehrenden umstritten geblieben, weil sie den Hochschulen eher aufgedrängt, als daß sie als eigenes Anliegen von ihnen betrieben wurde. Auslöser waren Hochschulrankings Anfang der 90er Jahre unter der Frage: „Welche Uni ist die beste?“ (SPIEGEL 1990 und 1993) oder Initiativen studentischer Verbände, zum Beispiel unter dem Motto oder „Mehr Ehre für die Lehre“ (RCDS 1992).

Was unter „Lehrevaluation“ gefaßt wird, ist aber sehr unterschiedlich in Verfahren wie Zielsetzung. In einem allgemeinen Sinne kann sie als kommunikativer Vorgang über die Lehre und deren Qualität an Hochschulen verstanden werden. Diese Verständigung über die Lehre kann sich auf verschiedene Quellen stützen. Kern sind zumeist studentische Stellungnahmen, die über Befragungen gewonnen werden: das Modell der studentischen Veranstaltungskritik. Sie sind auch Grundlage des zweitens Modells: dem Hochschulranking, wobei weniger die „Rückmeldung“ an Lehrende im Vordergrund steht, sondern der öffentliche Vergleich von Hochschulen und Fächern. Als drittes Modell hat sich der „Lehrbericht“ etabliert, der durch Landesgesetze den Hochschulen und Fakultäten abverlangt wird. Schließlich wird immer häufiger die gutachterliche Bestandsaufnahme mit Empfehlungen durch Fachexperten und Kollegen durchgeführt, das Modell der „Peer-Reviews“.

Mit diesen Modellen der Evaluation von Hochschullehre und ihren Verfahren der Datengewinnung wie -verwendung sind eine Reihe von Problemen verknüpft. Nachdem die vier Modelle vorgestellt worden sind (Abschnitt 1), wird auf einige Fragen, die sich als zentral für die Akzeptanz und Produktivität von Evaluationen erwiesen haben, eingegangen. Können studentische Beurteilungen eine verlässliche Grundlage der Evaluation sein? Wie sind die Instrumentarien zu gestalten, damit die gewonnenen „Daten“ nützlich sind? (Abschnitt 2). Wie läßt sich Lehrqualität oder Lehrleistung bestimmen und wie sind die Dimensionen, die Kriterien und Indikatoren dafür zu ermitteln? (Abschnitt 3). Welche Anwendungen von Evaluationen hinsichtlich des Ziels einer Qualitätsverbesserung der Lehre haben sich als vorteilhaft herausgestellt und welche Voraussetzungen sollten erfüllt sein? (Abschnitt 4).

Zu diesen Fragen werden oftmals Mutmaßungen angestellt. Sie sind einerseits mit den Befunden empirischer Untersuchungen zu konfrontieren, andererseits sind vorhandene Erfahrungen heranzuziehen, um darüber zu klären, inwieweit sie zutreffend oder unberechtigt sind.

1. Modelle und Verfahren der Lehrevaluation

Im Laufe der letzten Jahre haben sich vier grundsätzliche Modelle der Lehrevaluation an den Hochschulen herausgebildet und finden ihre Anwendung. Ihre Grundzüge und gewissen Varianten werden kurz dargestellt, die vorhandenen Erfahrungen damit zusammengefaßt sowie aufgetretene Schwierigkeiten oder Probleme benannt.

1.1 Studentische Veranstaltungskritik

Die Beurteilungen von Lehrveranstaltungen durch die Studierenden sind Ausgang und Kern der Evaluation der Lehre, insofern sie, über Befragungen ermittelt, Grundlage oder Teil – neben anderen Datenquellen – der weiteren Modelle und Verfahren sind. Ihre Angaben werden zumeist als direkte Rückmeldung für die Lehrenden genutzt. Sie bilden aber auch die Grundlage der Hochschulrankings und können als Teil der Lehrberichte oder bei den Peer-Reviews berücksichtigt werden.

Die studentischen Stellungnahmen und Urteile können sich auf eine einzelne Veranstaltung eines Dozenten beziehen. Häufiger umfassen sie eine ganze Veranstaltungsreihe, etwa von einem Semester, und werden gegen deren Ende erhoben. Zu beachten ist dabei die folgende Differenz: Ob die „Lehrleistung des Dozierenden“, die „Qualität der Veranstaltung“ oder der „Lernfortschritt der Studierenden“ Gegenstand der Evaluation ist – was oftmals vermischt wird und wogegen sich zu Recht manche Kritik richtet (vgl. klärend HELMKE 1996).

Darüber hinaus werden Studierende auch zur Lehr- und Studienqualität eines Fachbereichs bzw. Studienganges befragt. Bei dieser Bezugsebene sind erheblich andere Aspekte in den Blick zu rücken, als wenn nur eine einzelne Veranstaltung oder ein einzelner Lehrender evaluiert werden (vgl. WEBLER/DOMEYER/SCHIEBEL 1993). Dazu zählen die Abstimmung der Veranstaltungen und der Aufbau der Veranstaltungsfolgen, der Forschungs- und Praxisbezug im Studium, das soziale Klima im Fachbereich sowie der Aufbau und die Durchführung von Prüfungen – Themen übrigens, die den Studierenden oftmals wichtiger sind als die mehr oder weniger gelungene Veranstaltung einzelner Dozenten (vgl. BARGEL 1996).

Inwieweit schriftliche Fragebogen mit vorgegebenen Antwortkategorien für die Lehrevaluation durch Studierende einsetzbar sind, ist eine entscheidende und umstrittene Frage. Ihr Einsatz hängt davon ab, ob die Studierenden in der Lage sind, zutreffende Urteile abzugeben. Daß solche Befragungen verlässliche und gültige Resultate über die „Lehrleistung der Lehrenden“ oder die „Studienqualität eines Fachbereiches“ ermöglichen, sieht sich manchen Einwänden ausgesetzt (vgl. dazu ausführlicher Abschnitt 2).

Drei kritische Zonen zeigen sich bei der bisherigen Anwendung studentischer Befragungsdaten im Rahmen von Evaluationsvorhaben, um darüber eine Verbesserung der Lehre zu erreichen: (1) Die Inhalte und Items der Befragungsinstrumente sind oft unausgewogen und zu pauschal formuliert. Sie liefern damit kaum konkrete Hinweise, worauf sich mögliche Defizite beziehen und wie sie

abzustellen wären. (2) Die Auswertung und Darstellung der Resultate der Befragungen begnügen sich zu häufig mit den bloßen Rohdaten oder einfachen Mittelwerten; es werden zu selten Verteilungen, Profile und Zusammenhänge aufgezeigt. (3) Schließlich bleibt die Verwendung der Befunde durch die Lehrenden offen: Werden Folgerungen gezogen oder entstehen „Datenfriedhöfe“? Mit der Rückmeldung an die Lehrenden ist es nicht getan, denn es sollten von ihnen Konsequenzen gezogen und diese mit den Studierenden diskutiert werden. In manchen Fällen sollten Lehrende Hilfestellungen erhalten, wie sie ihre Veranstaltung in Anlage und Durchführung verbessern können.

Aufgrund der bisherigen Anwendungen des Instrumentes der studentischen Veranstaltungskritik (oft nur pauschale, sehr kurze Fragebogen, einfache Auswertungen über Mittelwerte, keine Zielkonzeptionen) werden die vorliegenden Erfahrungen oftmals dahingehend bilanziert, es seien zwar in Folge der Befragungen noch keine konkreten Verbesserungen in der Lehre nachweisbar, aber die Transparenz über die Lehrsituation habe sich erhöht und eine Diskussion über die Lehre und ihre Qualität sei wenigstens angestoßen worden (z.B. STANGL u.a. 1995).

Daß studentische Beurteilungen nicht immer auf aufwendigen Fragebogen beruhen müssen, belegen PETER/WAWRZINEK (1995) mit dem Verfahren der „Dialogischen Evaluation“. Sie hat zum Ziel, Lehre nicht nur zu bewerten, sondern unmittelbar Verbesserungen zu stimulieren. Bei der „Dialogischen Evaluation“ wird eine Lehrveranstaltung von studentischen Evaluatoren (die keine Kursteilnehmer sind) besucht. Nachdem die Studierenden mit Hilfe von Fragebögen eine Einschätzung der Veranstaltung abgegeben haben, werden in Kleingruppen die Probleme und Verbesserungsmöglichkeiten zusammengetragen, um anschließend zwischen dem Lehrenden und den Studierenden diskutiert zu werden.

Dieses Verfahren zeigt, daß mit recht einfachen Mitteln aus einer bloßen „Feedback-Runde“ zum Veranstaltungsende eine fast systematische, zielgerichtete Evaluation gestaltet werden kann. Für das Ziel einer Kommunikation über die Qualität der Veranstaltung erscheint dieses Verfahren besser geeignet als ein aufwendiger Fragebogen, der bloß der Rückmeldung dient. Jedoch lebt diese Art der Evaluation von dem Engagement der Studierenden und verlangt einigen organisatorischen Aufwand, um sie systematischer einzusetzen. Ähnliche studentische Diskussionsrunden werden mittlerweile auch zur Datensammlung bei der Beurteilung von Studiengängen bzw. Fachbereichen eingesetzt (BÜLOW-SCHRAMM 1995).

1.2 Rankings von Hochschulen und Fächern

Das Thema „Lehrqualität“ wurde in Deutschland erst vor gut zehn Jahren durch ein Ranking populärer. Der SPIEGEL (1989) ließ 6000 Studierende (12 pro Fach und Hochschule) ihren Fachbereich beurteilen. Die Mittelwerte dieser Daten bildeten die Grundlage für das erste hochschulvergleichende Ranking in Deutschland. Das Ergebnis lautete damals: Die neuen Hochschulen sind die besten. Zehn Jahre später hatten jedoch die ostdeutschen Hochschulen diesen den Rang abgelaufen (SPIEGEL 1999).

Anders als z.B. in Großbritannien (TIMES 1999) scheinen Rankings als Verfahren der vergleichenden Qualitätsbeurteilung an deutschen Hochschulen inhaltlich weniger akzeptiert zu sein, obwohl sie viel Aufregung verursachen. Die Ablehnung könnte damit zusammenhängen, daß hierzulande eine Konkurrenz zwischen den einzelnen Hochschulen weniger erwünscht ist, und Hochschulrankings nicht von Hochschulen, sondern von auflagestarken Magazinen in Auftrag gegeben werden (z.B. MANAGER-MAGAZIN 1994; FOCUS 1997).

Die methodische Qualität dieser Rankings ist heftig umstritten. So wurde z.B. an der zweiten Befragung des SPIEGEL (1993) bemängelt, daß die erstellten Ranglisten mangelhafte Produkte seien, die auf zufälligen Antwortmustern beruhen und höchstens journalistischen Wert haben (HORNBOSTEL/DANIEL 1994). Die Stichprobe der befragten Studierenden sei zu klein, willkürlich und nicht repräsentativ. Auch die Art der Indikatoren und ihre Gewichtung wird häufiger als unzureichend kritisiert.

Neben diesen methodischen Schwächen, denen spätere Rankings durch größere Fallzahlen und eine differenziertere Auswertung zu begegnen versuchen, bleibt jedoch für allgemeine Rankings die Frage offen, inwieweit Bewertungen zur Lehrleistung über eine gesamte Hochschule aussagekräftig und von praktischem Nutzen sein können. So kommt die Untersuchung von GIESEN und JANSEN (1983) zu dem Ergebnis, daß der Varianzanteil, der auf die spezifischen Erfahrungen in den Fachgruppen zurückgeführt werden kann, im Bereich der Lehrqualität stets größer ist als der universitätsspezifische Anteil.

Bei der Beurteilung der Lehrsituation und der Lehrqualität werden große Fachunterschiede an der jeweiligen Hochschule sichtbar. Die Fächer unterscheiden sich in der materiellen Ausstattung, der Betreuungsrelation (Lehrende/Studierende), bei den von den Studierenden eingeschätzten didaktischen Fähigkeiten der Lehrenden oder deren Zugänglichkeit für die Beratung der Studierenden. Diese Differenzen sind zwischen den einzelnen Fächern so gravierend, daß von unterschiedlichen Fachkulturen gesprochen werden kann. In den meisten Fällen divergieren zwei Fächer an einer Hochschule in vielen Merkmalen stärker voneinander als dasselbe Fach an verschiedenen Universitäten oder Fachhochschulen (BARGEL 1993).

Gegen vergleichende Rankings auf der Ebene von Studiengängen spricht dies – soweit sie den üblichen Gütekriterien entsprechen – jedoch nicht. Ganz im Gegenteil: Die Veröffentlichung der ersten Untersuchungsergebnisse zur Lehrqualität sorgte für viel öffentliche Aufmerksamkeit, so daß in der Folge an den Hochschulen im Bereich der Lehrevaluation einiges in Angriff genommen wurde. Gerade die übersichtliche Ergebnisdarstellung im Rankingverfahren erfüllt zwei Kriterien: Zum einen wird die Lehrqualität an den Hochschulen transparenter und öffentlicher, und zum anderen erzeugen derartige Vergleiche sozialen Druck, um zu Verbesserungen in der Lehrqualität zu motivieren. Sie vermitteln jedoch keine oder nur wenig Informationen, worin die Defizite oder Probleme in der Lehre bestehen. Auch die Vorgabe, danach könnten Studierende sich den „besten“ Hochschulort aussuchen, scheint angesichts der Differenz der Befunde im Vergleich der verschiedenen Rankings fragwürdig.

1.3 Lehrberichte der Hochschulen und Fachbereiche

Seit Anfang der 90er Jahre werden in Deutschland Lehrberichte als Mittel zur Beurteilung der Lehrqualität diskutiert und eingesetzt. Sie sind meist als vorwiegend statistische Datensammlungen angelegt, die eine Hochschule bzw. ein Studiengang zu Bereichen wie Fachstudiendauer, Schwund- und Erfolgsquote, Notenverteilung u.ä. erheben sollte. Die Erstellung von Lehrberichten wurde seit Mitte der 90er Jahre von seiten der Bundesländer auf eine gesetzliche Grundlage gestellt. In den verschiedenen Wissenschaftsgesetzen sind Lehrberichte meist als jährlicher bzw. zweijährlicher Bericht ohne genaue Definitionen vorgesehen. Als Ziel wird die Transparenz nach außen und die Bestandsaufnahme nach innen genannt.

Als weitere Ziele von Lehrberichten nennt WEBLER (1995) die Möglichkeit für einzelne Lehrende, Fachbereiche und Hochschulen, Mißstände und Fehlentwicklungen zu erkennen (Selbstkontrolle) und rationale, informiertere (Ressourcen-)Entscheidungen in Fachbereichen, Hochschulen und Ministerien treffen zu können, sowie für Studierende, daß sie dadurch mit wichtigen Informationen für Orts- und Fachwahlentscheidungen ausgestattet werden.

Das „Bielefelder Modell“ zur Erstellung von Lehrberichten (WEBLER/DOMEYER/SCHIEBEL 1993) ist die detaillierteste Anleitung für entsprechende Dokumentationen. Dennoch bleibt auch in diesem Modell die genaue Konzeptentwicklung in der Hand der Hochschule. Hochschulen können und sollen ihre eigenen Kriterien zusammenstellen, welche Angaben sie bzw. die Fachbereiche in ihren Lehrbericht aufnehmen wollen. Damit soll den Hochschulgremien die Möglichkeit gegeben werden, ihre eigenen Vorstellungen einzubringen. Hinzu kommt, daß die Akzeptanz in den Hochschulen größer ist und der Widerstand geringer ausfällt, wenn exakte und allgemein verbindliche Kriterien zur Erstellung entsprechender Dokumentationen nicht detailliert festgelegt sind.

Die starke Beteiligung der Hochschulen bei der Konzeption erscheint auch insoweit sinnvoll, da als oberstes Ziel der Lehrberichterstattung das „Monitoring“ gesehen wird, d.h. die Selbstinformation und Selbstevaluation. Als Folge dieser Informationsaufbereitung und Situationsanalyse wird erwartet, daß Kommunikations- und Einigungsprozesse erzeugt werden, die als Konsequenz einen Maßnahmenkatalog zur Verbesserung der Lehrsituation im Fachbereich haben. Lehrberichte wären somit Instrumente zur Organisationsentwicklung an Hochschulen. Um dies jedoch erreichen zu können, sind weitaus mehr Angaben als einige statistische Kennwerte erforderlich. Daher empfehlen WEBLER/DOMEYER/SCHIEBEL (1993) bei der Erstellung des Lehrberichts eine Reihe weiterer Informationen aufzunehmen: die Festlegung der Ziele der Ausbildung, ein Leitbild für die Hochschule bzw. die Fachbereiche, Erhebungen zur Sachausstattung und Auslastung, die Ermittlung der Beratungsangebote für Studierende sowie eine Erhebung über die Maßnahmen zur Qualitätssicherung in der Lehre. Zusätzlich sollten bei ausführlichen Berichten die Überprüfung der Attraktivität des Studiengangs und eine Verbleibstatistik der Studierenden hinzukommen.

Für die Erhebungen sind Fragebögen für Lehrende (Meinungen und Auffassungen zur Lehre) und für Studierende (zur Situation im Studiengang/Fach-

bereich) vorgesehen. Des weiteren können Interviews mit zentralen Personen (Dekan, Fachschaft etc.) oder eine Dokumentenanalyse fachbereichsspezifischer Schriftstücke (z.B. Prüfungsordnung, Protokolle des Fachbereichsrats) einbezogen werden. Bezugspunkt zur Dateninterpretation können Fachbereiche der gleichen Fachrichtung an anderen Hochschulen sein oder Zeitreihenanalysen desselben Fachbereichs.

Die Anforderung, wonach Lehrberichte neben Daten der allgemeinen Hochschulstatistik auch qualitative Ergebnisse und angestrebte Verbesserungsmaßnahmen umfassen sollten, konnte sich bisher nicht durchsetzen. Nach den in den letzten Jahren gesammelten Erfahrungen umfassen Lehrberichte häufig keine Evaluationsdaten oder Ergebnisse studentischer Veranstaltungskritik (WOLTER 1996). Meist sind nur Daten der „quantifizierbaren Effizienzkontrolle“ enthalten. Diese Einschätzung hängt damit zusammen, daß Hochschulen es immer noch vermeiden, Kriterien „guter Lehre“ zu formulieren. In den gesetzlichen Grundlagen sind meist keine ausreichenden Kriterien enthalten, nach denen Lehrberichte zu erstellen sind (vgl. die Übersicht bei HAGE 1996a). Die Berichte einzelner Fachbereiche sind folglich häufig sogar innerhalb einer Hochschule nicht einheitlich. Diese Arten von Datensammlungen können nach SÖLLNER (1996) eher als „Improvisationen“ eingestuft werden.

Die Wirksamkeit der Lehrberichte und ihr Stellenwert für Verbesserungen in der Lehre sind bislang gering geblieben. Die weitgehende Wirkungslosigkeit, die auf verschiedenen Tagungen diskutiert wurde, hat auch eine Befragung von Dekanen in Nordrhein-Westfalen bestätigt (MINISTERIUM FÜR WISSENSCHAFT UND FORSCHUNG 1998). Damit Lehrberichte zur Entwicklung der Lehrqualität beitragen, so die Folgerung, sollte eine zentrale Instanz an der Hochschule die Erstellung der Lehrberichte koordinieren, eine Auswertung der Lehrberichte muß erfolgen und eine inhaltliche Rückkoppelung der Auswertungsergebnisse an die betreffenden Fachbereiche sollte institutionalisiert sein.

Lehrberichte dienen in erster Linie zur Selbstevaluation, um an den Hochschulen im jeweiligen Fachbereich die Situation der Lehre transparenter zu machen. Sie sind meist auch die erste Stufe des Verfahrens der Peer-Evaluation, bei dem Fachkollegen die erhobenen Daten sichten, teilweise eigene Erhebungen durchführen und Maßnahmen zur Verbesserung der Lehrqualität formulieren.

1.4 Peer-Reviews und kollegiale Begutachtungen

In Deutschland haben sich seit Beginn der 90er Jahre zur Beurteilung von Fachbereichen Vorhaben etabliert, die nach dem sogenannten „Niederländischen Modell“ der Lehrevaluation vorgehen. Nach einer internen Datensammlung zur Lehr- und Studiensituation an einem Fachbereich werden Fachkollegen aus anderen Hochschulen (sog. „Peers“) ausgewählt, die vor Ort eine Begutachtung der zusammengetragenen Daten durchführen, Interviews mit den Beteiligten arrangieren und einen Bericht verfassen. Dieses Verfahren ist daher unter dem Begriff „Peer-Review“ bekannt geworden (HOLTKAMP/SCHNITZER 1992; REISSERT/CARSTENSEN 1998).

Nach der Erhebung des Ist-Zustands werden von den Peers in einem weiteren Schritt konkrete Empfehlungen formuliert. Das Modell sieht nach einer gewissen Zeit ein Bewertungstreffen vor, bei dem die Umsetzungen der vorgeschlagenen Maßnahmen besprochen werden sollen. Inzwischen haben sich in einigen Bundesländern organisatorische Einheiten gebildet, die Auswahl und Durchführung der Peer-Reviews koordinieren (z.B. Niedersächsische Evaluationsagentur).

Die Erfahrungen in den Niederlanden, in denen bereits einige beurteilende Berichte zur Peer-Review an Hochschulen vorliegen (vgl. RICHTER 1996), zeigen eine hohe Akzeptanz des Verfahrens an den beteiligten Hochschulen. Dies ist nicht zuletzt dadurch bedingt, daß Fachkollegen eine ausreichende Sachkenntnis zugebilligt wird und die selbständige Auswahl der Beurteiler eine Vertrauensbasis herstellt.

Die Probleme der Peer-Reviews sind vergleichbar mit den Schwierigkeiten, die bei der Erstellung von Lehrberichten auftreten. Notwendige statistische Daten, die Studienverläufe beschreiben, werden von den Fachbereichen häufig nicht erhoben, unzureichend ausgewertet oder sind von schlechter Qualität. Erstaunlicher ist dagegen die Erfahrung, daß bei der Peer-Review inhaltliche Aspekte des Studiums im Rahmen der Beurteilung vernachlässigt werden. Dabei ist eines der Hauptargumente zur Bevorzugung von Kollegen im Vergleich zu Studierenden für die Bewertung der Lehre, daß erstere für die inhaltliche Einschätzung des Studiums notwendige Sachkenntnisse besitzen.

Möglicherweise werden kaum Aussagen zu den Studieninhalten getroffen, weil kritische Punkte – über alle Bereiche hinweg – nicht ausreichend angesprochen werden. Hinzu kommt, daß viele Empfehlungen zu allgemein formuliert sind und häufig nicht überprüft wird, inwieweit die Verbesserungsvorschläge der Peers tatsächlich umgesetzt wurden. Letzteres dürfte darauf zurückzuführen sein, daß keine Sanktionen für die ausbleibende Umsetzung der Empfehlungen vorgesehen sind.

Auch in Deutschland konnten erste Erfahrungen zu Ergebnissen und Schwierigkeiten von Evaluationsmaßnahmen gesammelt werden (z.B. REUKE 1996). Ein Problem sind auch hier oftmals fehlende Daten in der Selbstevaluation der Fachbereiche. Diese Lücken sind vor allem da zu finden, wo die didaktischen Fähigkeiten der Lehrenden eingeschätzt werden sollen sowie bei der Beurteilung zum Praxisbezug der Lehre.

Der heikelste Punkt an den Hochschulen ist die Frage der Veröffentlichung bzw. Nichtveröffentlichung von Evaluationsergebnissen. Obwohl sich der WISSENSCHAFTSRAT (1996) in seinen Empfehlungen uneingeschränkt für eine möglichst breite Öffentlichkeit ausspricht, befürworten nur wenige Fachbereiche diese Variante. REUKE (1996) berichtet, daß die Ergebnisse der Peer-Evaluation in den einzelnen Hochschulgremien durchaus wahrgenommen werden, die Umsetzung der Empfehlungen jedoch nur bedingt erfolgt. Insgesamt kann die Folgerung gezogen werden, daß die Ergebnisse der Peer-Reviews, ähnlich wie die der Lehrberichte, das Thema „Qualität der Lehre“ stärker in die Diskussion an den Hochschulen gebracht haben. Der Schritt zu einer systematischen Verbesserung aufgrund dieser Verfahren wurde jedoch noch nicht vollzogen. Alles in allem erscheint dieses Verfahren aber eher als das der bloßen Lehrberichte ein Weg zur Verbesserung der Lehre zu sein.

2. Probleme studentischer Lehrbeurteilungen

Studentische Beurteilungen zur Studien- und Lehrqualität bieten leicht zugängliche und ökonomisch günstige Datengrundlagen. Allerdings wird angezweifelt, ob die Studierenden in der Lage sind, zutreffende Urteile über die Lehre oder die Lehrleistung der Professoren abzugeben. Deshalb könnten studentische Befragungen keine gültigen und verlässlichen Resultate erbringen, an denen sich die Entwicklung der Lehre orientieren sollte (vgl. die Zusammenstellung solcher Argumente bei ALEAMONI 1987; im Überblick auch in HAGE 1996a).

Die „mangelnde Sachkenntnis“ der Studierenden ist wohl das häufigste Argument gegen eine Beurteilung von Lehrveranstaltungen durch Kursteilnehmer. Studierende würden nicht den inhaltlichen Wert einer Veranstaltung oder die tatsächliche Lehrleistung der Dozierenden einschätzen, sondern würden vielmehr nur ihren Unterhaltungswert oder die bloße „Beliebtheit“ des Lehrenden bewerten. Im folgenden wird deshalb etwas ausführlicher auf dieses Argument gegen den Einsatz studentischer Veranstaltungskritik eingegangen, das immer wieder thematisiert wird (z.B. KROMREY 1996, S. 69).

2.1 Inhalts- oder Unterhaltungswert

In einer Anzahl von Studien, als Dr. Fox-Studien bekannt geworden, wurde untersucht, ob Studierende eine Veranstaltung vor allem nach ihrem Unterhaltungswert beurteilen. Die ersten dieser sozialen Experimente belegten, daß Studierende denselben Vortrag eines angeblichen Dozenten (Dr. Fox, der von einem Schauspieler dargestellt wurde) bei höherem Unterhaltungswert besser beurteilten (vgl. NAFTULIN/WARE/DONNELLY 1973). Das sprach für die Annahme, Studierende orientierten sich hauptsächlich am unterhaltsamen Vortrag und weniger an der inhaltlichen Qualität, wenn sie die Lehrleistungen eines Dozenten beurteilen.

WARE/WILLIAMS (1975) überprüften in einer zweiten Reihe von Experimenten diese Ergebnisse. Sie stellten fest, daß Studierende den Inhalt eines Vortrages später besser wiedergeben konnten, falls der Vortragsstil expressiver war. Der in beiden Fällen identische Dozent erhielt bei einem lebendigeren Vortrag außerdem eine bessere Beurteilung seiner Lehrleistung. Nach diesen Befunden erscheint, angesichts der höheren Behaltensleistung der Studierenden, eine bessere Bewertung der Lehrleistung aufgrund eines anregenden Vortrages durchaus gerechtfertigt.

In einer dritten Version der Studie kontrollierten MARSH/WARE (1982) erneut den „Dr. Fox“-Effekt. Ausgehend von der These, daß Studierende mit einer bestimmten Zielmotivation Lehrveranstaltungen besuchen (z.B. mit Blick auf anschließende Prüfungen), wurde versucht, die extrinsische Motivation der Wissensaufnahme zu variieren. Diese differenzierte Untersuchung kam zu dem Ergebnis, daß ein lebendiger Vortragsstil vor allem dann wichtig ist, wenn die extrinsische Motivation niedrig ist, d.h. die Studierenden das vermittelte Wissen nicht zielgerichtet benötigen. Wenn allerdings die Studierenden bereits motiviert sind, z.B. weil der Inhalt der Veranstaltung in einer Prüfung abgefragt

wird, erachten sie die Inhaltsdichte im Vergleich zum Stil des Vortrags als relevanter für ihre Einschätzung der Qualität der Veranstaltung.

Die Ergebnisse dieser Studie belegen eindrucklich, daß Studierende zwischen dem Unterhaltungswert und der inhaltlichen Qualität einer Veranstaltung unterscheiden, und sie den inhaltlichen Wert eines Vortrags in Hinblick auf universitäre Anforderungen höher gewichten. Die Sorge, die Ausrichtung der Studierenden am Unterhaltungswert einer Veranstaltung würde die Beurteilung der Lehrleistung verzerren, erweist sich deshalb als unbegründet.

2.2 Argumente gegen studentische Veranstaltungskritiken

Einwände, studentische Urteile zur Lehre seien unzutreffend, führen eine Vielzahl möglicher „Verdachtsmomente“ an. Eine Reihe solcher Variablen, die im Verdacht stehen, studentische Veranstaltungskritik zu verzerren, wurde in einer Vielzahl von Studien und mit verschiedenen Methoden überprüft. Eine ausführlichere Dokumentation entsprechender Ergebnisse ist bei HAGE (1996a) zu finden. Im folgenden sind einige Ergebnisse zusammengefaßt.

Geschlechtsspezifische Unterschiede in den Urteilen über die Studienverhältnisse oder von Lehrveranstaltungen konnten pauschal in keiner Studie ermittelt werden (ELMORE/POHLMANN 1978; HORNBOSTEL/DANIEL 1994; RINDERMANN 1996). Ebenso konnten keine signifikanten Unterschiede in der Lehrbewertung von Studierenden in unterschiedlichen Semestern gefunden werden (DANIELSEN & WHITE 1976; DANIEL/THOMA/BANDILLA 1994; RINDERMANN 1996).

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 (DANIEL 1995; HAGE 1996b). Die Ergebnisse zeigen ebenfalls, daß weder ein Hochschulwechsel noch ein Auslandsstudium oder die Mitarbeit an einem Forschungsprojekt zu anderen Urteilen der Studierenden führen (HAGE 1996b; HORN-BOSTEL/DANIEL 1994).

Bestätigt wurde jedoch die Annahme, daß ein unterschiedlich ausgeprägtes Interesse am Studium bzw. am entsprechenden Fach die Lehrbeurteilung beeinflusst oder gar determiniert (HOFMANN 1988; ESSER 1995). KROMREY (1993/1994) geht sogar so weit, die unterschiedlichen Lehrbewertungen vor allem auf das unterschiedliche Interesse der Studierenden zurückzuführen. Auch RINDERMANN (1996) fand beträchtliche Übereinstimmungen zwischen dem Vorinteresse der Studierenden am Kursthema und ihrer anschließenden Beurteilung des entsprechenden Kurses. DANIEL (1995) ermittelte, 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. Insgesamt kann daher als gut belegt gelten, daß ein höheres Interesse der Studierenden zu einer positiveren Einschätzung der Lehrqualität führt, aber deren Stärken und Schwächen gleichermaßen erkennbar bleiben.

Was ebenfalls die Beurteilung von Lehrveranstaltungen beeinträchtigen dürfte, ist der Faktor „Überfüllung“. Fachbereiche, in denen Vorlesungen,

Übungen oder Seminare häufig überfüllt sind, erfahren insgesamt schlechtere Beurteilungen hinsichtlich der Studien- und Lehrqualität. Vor allem die hochschuldidaktische Qualität der Durchführung von Lehrveranstaltungen und die tutoriale Qualität der Beratung und Betreuung durch die Lehrenden werden dann von den Studierenden deutlich schlechter eingeschätzt (BARGEL/RAMM/MULTRUS 1996; HAGE 1996b).

Um Lehrende aufgrund von Ergebnissen studentischer Veranstaltungskritik miteinander vergleichen zu können, muß der Einfluß der jeweiligen Veranstaltung (Typ, Größe, Relevanz) möglichst niedrig sein. Oft wird dazu die These vertreten, daß nicht die Lehrenden, sondern der Kurs die Lehrqualität, gemessen anhand der Urteile der Studierenden, bestimmt. Die Ergebnisse einer umfassenden Untersuchung von MARSH (1982) weisen jedoch nach, daß zwar gewisse Varianzanteile existieren, die nur auf den entsprechenden Kurs zurückzuführen sind, daß jedoch der Anteil der konsistenten Beurteilung der Lehrenden im Vergleich dazu deutlich höher liegt und für eine vergleichende Lehrbewertung ausreichend erscheint.

Den gesichteten Studien zufolge ist es möglich, über Fragebogen der studentischen Veranstaltungskritik Lehre und Lehrende zutreffend und unverzerrt beurteilen zu lassen, auch für Zwecke des Vergleichs zwischen Lehrenden oder Fachbereichen. Allerdings ist zu bedenken: Studierende sind unterschiedlich interessiert und leistungsfähig, Fachbereiche oder Veranstaltungen sind aufgrund ihrer Größe und Zusammensetzung der Studierenden nur bedingt vergleichbar. Sollen daher Vergleiche angemessen und fair sein, sind die Voraussetzungen der Studierenden (Interesse und Leistungsfähigkeit) ebenso wie die Ausgangslage in den Fächern (Auslastung und Ausstattung) zu berücksichtigen. Diese sollten entweder kontrolliert oder auspartialisiert werden.

3. Dimensionen und Instrumente der Lehrevaluation

Die Ansprüche an die Instrumente, anhand derer die Studierenden eine Veranstaltung beurteilen, sind abhängig von ihrer Verwendung: Werden sie nur als interne Rückmeldung für die Lehrenden gebraucht, die daraus ihre eigenen Schlüsse ziehen, oder dienen sie als Grundlage öffentlicher Vergleiche der „Lehrleistung“ von Dozierenden, anhand derer Studierende Hochschulen oder Fächer wählen sollen, oder fungieren sie gar als Mittel für politische Steuerungsverfahren (z.B. für Gehalt und Karriere der Lehrenden, Mittelvergabe an Fachbereiche)? Je weitreichender die Verwendung angelegt ist, desto strenger müssen die methodischen Anforderungen an die Instrumente sein.

Wenig transparente und scheinbar verzerrte Ergebnisse von Lehrbewertungen liegen oftmals an einer mangelnden Qualität der verwendeten Fragebogen. Eine angemessene Erfassung von Lehrleistungen durch studentische Stellungnahmen ist daher mit mehrdimensionalen Fragebogen empfehlenswert (MARSH 1982, AMELANG/RINDERMANN 1994). Je genauer ein entsprechender Bogen ausgearbeitet ist, desto geringer ist die Gefahr, daß externe Faktoren die Ergebnisse verzerren.

Bezogen auf CHRISTIAN (1974) muß aber festgestellt werden, daß ein deduktiv-theoretisches Herangehen bei der Ermittlung von Lehrdimensionen bislang

kaum möglich ist, da eine „Theorie des Unterrichts“ an der Hochschule (noch) nicht existiert (bereits SCHOTT 1973). In den vergangenen Jahren hat sich in diesem Punkt zwar einiges entwickelt, so wurde z.B. die kognitive Informationsverarbeitung aufgegriffen und im schulischen Lernen untersucht. Dennoch ist zum Beispiel die Pädagogische Psychologie von einer solchen „Theorie des Unterrichts“ noch weit entfernt. WEINERT (1986) führt diesen Mangel auf die „disfunktionale Arbeitsteilung“ zwischen dem Auffinden elementarer psychischer Lernmechanismen und der angewandten pädagogischen Forschung zurück. Eine weitere Ursache für die fehlende theoretische Grundlage im Bereich der Lehre spricht HECKHAUSEN (1986) an: die Bindung der Pädagogischen Psychologie an die Institution Schule und die üblichen Schüler-Altersstufen. Theoretische Fortschritte in dem komplexen Problemfeld von Lehren und Lernen an der Hochschule sind daher für die Gestaltung der Lehre wie deren Evaluation von großer Wichtigkeit (vgl. z.B. EITEL 1995). Im Gegenzug können Vorhaben der Lehrevaluation, vorausgesetzt sie sind nicht nur pragmatisch und ad hoc ausgelegt, sondern in theoretische Überlegungen und Annahmen eingebettet, durchaus die Lehr-Lern-Forschung befördern (vgl. HELMKE 1996). Freilich gehen sie dann über den Status einfacher Rückmeldungen an Lehrende hinaus und richten ihr Augenmerk auf Dimensionen, Strukturen und Zusammenhänge.

In den bisher durchgeführten Untersuchungen zur Ermittlung von Dimensionen der Hochschullehre liegt überwiegend eine induktive oder pragmatische Herangehensweise vor. Verwendet werden zum Beispiel die „Critical Incidents Technique“, „manuelle“ Klusterung von Dozentenbeschreibungen zu homogenen Kategorien, multidimensionale Skalierungen und ähnliches. Es wurden auch Sekundäranalysen bereits ermittelter Faktoren durchgeführt und diese zu übergreifenden Dimensionen zusammengeführt. Die nachfolgenden acht Dimensionen der Lehrbewertung sind in verschiedenen Studien immer wieder anzutreffen (BEATTY/ZAHN 1978). Sie sollten als „Grundraster“ für die Evaluation von Lehrveranstaltungen berücksichtigt werden, um die Lehrleistung differenziert und angemessen zu erfassen.

Die Dimension „Zuwendung“ umfaßt vor allem Aspekte der Freundlichkeit und Aufgeschlossenheit der Lehrenden, die Sensibilität für die Reaktion der Studierenden, die Toleranz für andere Blickwinkel, die Zugänglichkeit für Kritik, Respekt und Ermutigung gegenüber den Studierenden. Die Erreichbarkeit der Lehrenden und ihre Bereitschaft zur Beratung der Studierenden, nicht nur auf die formale Sprechstunde begrenzt, sind ebenfalls Aspekte dieser Dimension.

Der Dimension „Fairneß“ von Prüfungen und Benotungen ist wegen der großen Rolle, die Examen und Noten an den Hochschulen und für die Studierenden spielen, besondere Aufmerksamkeit zuzuwenden. Sie wird in vielen Evaluationsinstrumenten übersehen, obwohl faire und gerechte Leistungsrückmeldungen oder Prüfungen für die Studierenden sehr wichtig sind. Gelegentlich werden dazu auch Fragen zur Prüfungsrelevanz des Stoffes aufgenommen, die als Indikatoren für die Fairneß von Prüfungen angesehen werden.

Die Dimension „Kommunikationsfähigkeit“ bezieht sich auf die didaktische Fähigkeit der Lehrenden, abstrakte und komplexe Ideen und Theorien darstellen zu können, verständlich und gut zu erklären und auf die Voraussetzungsadäquatheit der Stoffvermittlung zu achten. Hinzu kommen Fragen nach dem

Tempo und der Prägnanz der Wissensvermittlung, der Verwendung von Hilfsmitteln bis hin zur akustischen Verständlichkeit.

Die Dimension „Kurs- bzw. Stofforganisation“ beinhaltet üblicherweise Items zu den thematischen Schwerpunkten und ihrem Zusammenhang in einem Kurs, zur Systematik und Zeitökonomie sowie zur Koordination und Vororientierung einer Veranstaltung. Damit sind wichtige Voraussetzungen für die Bewältigung der Anforderungen durch die Studierenden angesprochen.

Im Unterschied zu dieser eher organisatorischen Ebene zielt die Dimension „Stimulierung“ stärker auf motivationale Aspekte ab. Dabei wird die Fähigkeit der Lehrenden beurteilt, das Interesse und die Eigeninitiative der Studierenden zu wecken sowie allgemein begeistern zu können.

Teilweise wird letzteres in einer eigenen Dimension „Enthusiasmus“ zusammengefaßt, worunter die Dynamik, das Charisma und die Anziehungskraft des Dozenten fällt. Obwohl die Beschreibung der Dimension „Enthusiasmus“ auf eine schwer veränderliche Persönlichkeitseigenschaft hindeutet, spielt dieser Faktor in der Lehre eine große Rolle (MURRAY 1983). Inwieweit diese Dimension jedoch Bestandteil studentischer Veranstaltungskritik sein sollte, ist umstritten und mag dahingestellt bleiben.

Sinnvoller erscheint es, stärker die Dimension „Variabilität vs. Monotonie“ zu berücksichtigen. Dabei handelt es sich um einen wichtigen Aspekt der Informationsaufnahme (SCHOTT 1973), der eine notwendige Voraussetzung dafür ist, für längere Zeit aufmerksam zu bleiben und den Wissensstoff zu rezipieren.

Die Dimension „Kurswert“ erfaßt schließlich den Kurs aus einer externen Perspektive. Ermittelt wird dabei die Bedeutung des Kursmaterials sowie dessen Prüfungsrelevanz und das Theorie-Praxis-Verhältnis. Im Vergleich zu den anderen Dimensionen wird weniger die Qualität des Lehrenden in den Mittelpunkt gestellt, als die Qualität und Relevanz des angebotenen Stoffes. In diesen Bereich werden auch Einschätzungen zur Kursschwierigkeit oder zur Arbeitsbelastung im Kurs ermittelt. Da belegt werden konnte, daß diese Aspekte kaum mit der Bewertung der Lehrleistung korrelieren, sollten sie nur sehr bedingt in die Beurteilung von Lehrenden einfließen.

Bei den Dimensionen der Lehrqualität und den damit zusammenhängenden Ergebnissen muß berücksichtigt werden, daß die meisten ermittelten Dimensionen das Resultat bestehender Fragebogen sind. Die dort aufgenommenen Items sind überwiegend intuitiv und pragmatisch formuliert und mit Hilfe von Faktorenanalysen oder ähnlichen statistischen Verfahren zu abgrenzbaren Dimensionen zusammengefaßt. Es wäre demnach möglich, daß die bisher verwandten Items relevante Elemente der Lehre vernachlässigen. Dies verdeutlicht die Notwendigkeit, pädagogische, psychologische und andere Theorien zur Formulierung von Kriterien guter Hochschullehre heranzuziehen. Erst ein Vergleich mit Items, die aus einer „Theorie der Hochschullehre“ abgeleitet sind, könnte beantworten, inwieweit sich die theoretisch abgeleiteten Dimensionen von den bisher intuitiv, pragmatisch und empirisch konstruierten unterscheiden würden.

Instrumente zur Evaluation von Lehrveranstaltungen können so gestaltet werden, daß sie Ansprüchen an die Validität und Reliabilität von Messungen im sozialwissenschaftlichen Bereich genügen und diese zum Teil sogar zu übertreffen vermögen. Deshalb ist ihre Verwendung im Rahmen der vergleichenden

den Lehrevaluation und Lehrberichterstattung vertretbar. In Deutschland existieren bereits einige elaborierte Instrumente der studentischen Veranstaltungskritik. Hierbei sind vor allem die Verfahren von DIEHL sowie von AMELANG und RINDERMANN zu nennen. DIEHL hat bereits in den 70er Jahren einen Fragebogen zur studentischen Veranstaltungskritik erarbeitet, der inzwischen als Version zur Beurteilung von Vorlesungen und als Version zur Beurteilung von Seminaren vorliegt (DIEHL 1994). Von praktischem Interesse ist dabei, daß das Anleitungshandbuch, die Fragebogen und das Auswertungsprogramm auf Diskette erhältlich sind und beliebig vervielfältigt werden dürfen. AMELANG und RINDERMANN haben das von ihnen entwickelte umfängliche Instrument (Heidelberger Inventar zur Lehrveranstaltungsevaluation, HILVE) ebenfalls publiziert, allerdings steht es nicht kostenlos zur Verfügung (AMELANG/RINDERMANN 1994).

Jedoch helfen auch die besten Instrumente nicht weiter, falls ihre Ergebnisse nicht sorgfältig ausgewertet und berichtet werden. Darauf weist nachdrücklich KROMREY (1994, 1996) wiederholt hin. In erster Linie ist dabei von Bedeutung, die Ergebnisse entsprechend ihrer Dimensionen und nicht nur pauschal auszuwerten. Neben den Mittelwerten sollten zudem die Verteilungen angegeben werden. Besonders aufschlußreich kann für den jeweiligen Lehrenden die Darstellung in Form von Profilen sein, der Ausweis von Stärken und Schwächen sowie das Aufzeigen von Zusammenhängen mit dem Kontext der Veranstaltung (vgl. dazu auch die Anregungen bei HELMKE 1996, S. 184).

4. Verbesserungen der Lehre mit Hilfe von Evaluationsdaten

Neben den Zweifeln, ob Studierende in der Lage sind, die Qualität einer Veranstaltung oder die Lehrsituation in einem Studiengang zu beurteilen, werden im Zusammenhang mit der Verwendung studentischer Bewertungen auch eine Reihe weiterer Fragen diskutiert. Sie berühren so wichtige Themen wie die Rolle des Lehrenden für den Studienerfolg, die Publikation von Evaluationsergebnissen und die Reichweite der Folgerungen, die auf ihrer Grundlage gezogen werden dürfen.

Bei der Frage, welche Stellung Lehrende in einer Lehrveranstaltung einnehmen, stoßen kontroverse Standpunkte aufeinander. Auf der einen Seite wird die Ansicht vertreten, daß es falsch sei von der „Qualität der Lehre“ zu sprechen, vielmehr ginge es um den „Erfolg des Studiums“ und dafür sei der „Eigenbeitrag der Studierenden“ von gleicher Bedeutung wie die „Lehrleistung der Dozenten“ (vgl. KROMREY 1996). Diese Position unterscheidet sich diametral von anderen Ansätzen, die Studierende als bloße „Kunden“ der Hochschulen sehen und diesen über die Lehrevaluation mehr Kundenorientierung abverlangen wollen. In der Lehre nehmen die Studierenden demnach eine „Dienstleistung“ in Anspruch, für die ihr Eigenbeitrag keine Rolle spielt.

Sicherlich kann festgehalten werden, daß der Erfolg einer Lehrveranstaltung nicht nur vom Engagement des Lehrenden, sondern auch von der Beteiligung der Studierenden abhängt. Inwieweit sich Studierende in Lehrveranstaltungen beteiligen und aktiv mitwirken können, wird durch verschiedene Faktoren bestimmt, wie z.B. die Anzahl der Teilnehmer, die Auslastung durch

andere Seminare, die Art der Veranstaltung. In der Verantwortlichkeit der Lehrenden bleibt jedoch, Möglichkeiten bereitzustellen, die eine Eigenbeteiligung der Studierenden eröffnen oder erhöhen. Insofern sollte dieser Aspekt selbst als ein Element der Lehrqualität thematisiert werden.

Neben solchen grundlegenden Fragen über Sinn und Anlage von Studium und Lehre wird im Zusammenhang mit dem Einsatz von Modellen und Instrumenten zur Lehrevaluation auch eine ganze Reihe praktischer Probleme diskutiert. Dabei handelt es sich um die Vorbereitung solcher Verfahren und ihre Akzeptanz durch die Dozierenden, die Wahl eines geeigneten Instruments, die sorgfältige Durchführung und Auswertung von Erhebungen sowie um die Frage, auf welche Art und Weise die Ergebnisse an die Lehrenden rückgemeldet werden und was sie letztlich bezwecken sollen.

Wird die Durchführung von Lehrevaluationen in einem Fachbereich oder einer Hochschule geplant, so ist eine sorgsame Einführung für die Betroffenen vorauszusetzen. Je transparenter und nachvollziehbarer das Verfahren vorgestellt wird, desto weniger ist mit Irritationen zu rechnen, die unnötige Probleme verursachen können (vgl. COMMELI 1991). Eine Möglichkeit, mit den Bedenken gegen die Beurteilung der Lehrqualität umzugehen, ist die Wahl von Instrumenten mit relativ hoher Meßgüte. Solche komplexen Meßverfahren haben allerdings den Nachteil, daß sie recht umfangreich sind, einen hohen Aufwand beim Einsatz wie auch in der Auswertung erfordern. Daher sollte der Verwendungszweck des Instruments bereits vor dem Einsatz geklärt sein. Es bestehen große Unterschiede bei den Anforderungen an Instrumente, je nachdem welchem Zweck sie dienen sollen. Es bedarf nur eines relativ einfachen Verfahrens, um einem Lehrenden ein kurzes Feedback zu geben. Soll ein Instrument dagegen Vergleiche zwischen Lehrenden ermöglichen oder sollen gar Ressourcen entsprechend der ermittelten Evaluationsergebnisse verteilt werden, sind methodische Gütekriterien (Zuverlässigkeit, Meßgenauigkeit, Objektivität) in viel stärkerem Maße zu beachten.

Zur Zeit ist die entscheidende Größe, die eine mögliche Auswirkung studentischer Evaluationsdaten bestimmt, die Motivation der Lehrenden. Die ermittelten Ergebnisse sollten daher differenziert ausgewertet und den Lehrenden ausführlich rückgemeldet werden (KROMREY 1996). Dennoch sollte bewußt bleiben, daß allein die Erhebung und Präsentation von Lehrbeurteilungen wenig Konsequenzen hat. CENTRA (1980) hat untersucht, unter welchen Bedingungen ein Feedback der Beurteilungsergebnisse Lehrende zur Veränderung der Unterrichtsgestaltung motiviert. Dabei kommt er zu dem Schluß, daß dies vor allem dann der Fall ist, wenn Lehrende in Lehraspekten, die sie selbst als wichtig einstufen und in denen sie eine weitaus bessere Selbsteinschätzung hatten, schlechtere Werte erzielten als von ihnen erwartet. Nur wenn Dozenten eine unrealistische Selbsteinschätzung in bezug auf die ihnen relevanten Lehraspekte hatten, sind sie demnach allein durch die Ergebnisse von Studienbefragungen zu Änderungen in ihrer Lehrpraxis bereit.

Zu ähnlichen Ergebnissen kommt RINDERMANN (1996), der bei einer wiederholten Lehrbefragung des gleichen Dozenten eine sehr hohe Stabilität der studentischen Urteile ermittelte. Das Ausbleiben von Veränderungen trotz Rückmeldung führt er auf zwei Ursachen zurück: Zum einen auf das Fehlen von Beratung und hochschuldidaktischer Weiterbildung und zum anderen auf das

Nichtvorhandensein eines Anreizsystems zur Verbesserung der Lehre. Es wäre daher eine wichtige Aufgabe, Angebote zur Aus- und Weiterbildung von Lehrenden als einen wichtigen Baustein zur Förderung der Qualität der Lehre an den Hochschulen zu entwickeln und zu etablieren (vgl. WINTELER/KRAPP 1999).

Die Zeit, bei der noch auf verschiedenen Tagungen (z.B. GRÜHN/GATTWINKEL 1992; LOCCUMER PROTOKOLLE 1992; BERENDT/STARY 1993) diskutiert wurde, ob und inwieweit die Hochschulen sich überhaupt mit der Qualität der Lehre und der Lehrevaluation beschäftigen sollten, ist offensichtlich vorbei. Inzwischen werden Fragebogen zur studentischen Veranstaltungskritik an verschiedenen Hochschulen eingesetzt. Dabei handelt es sich häufig um aufwendige einmalige Aktionen in einem gesamten Fachbereich oder um einen wiederholten, aber un-systematischen Einsatz in einzelnen Lehrveranstaltungen.

Studentische Lehrevaluation ist jedoch nur dann sinnvoll, wenn sie einen Nutzen hat, einen „Beitrag zur Optimierung der Lehre“ (RINDERMANN 1996, S. 93) leistet. Nur daran läßt sich festmachen, ob es sich um eine „Sackgasse“ handelt oder um ein Modell mit zukunftssträchtiger Perspektive. Zur Zeit ist allerdings noch offen, welchen Beitrag Lehrbeurteilungen von Studierenden in dieser Hinsicht tatsächlich leisten können.

Der deutlichste Effekt scheint zu sein, daß der Begriff der „Lehrqualität“ in weitaus mehr Hochschulgremien diskutiert wird als dies noch vor einigen Jahren der Fall war. Zum anderen sind jedoch diejenigen, die zu Beginn solche Befragungen oft in eigener Regie initiiert und durchgeführt hatten und die den meisten Nutzen davon haben sollten, die Studierenden nämlich, häufig enttäuscht, weil erkennbare Folgen ausbleiben.

Dies mag damit zusammenhängen, daß Zweck und Ziel studentischer Befragungen noch nicht vollständig geklärt sind. Lehrende sehen entsprechende Fragebogen in erster Linie als Instrument zum persönlichen Feedback, teilweise auch zur Kommunikation über die Lehre. Die Erwartung der Studierenden sind demgegenüber weitreichender. Sie erwarten erkennbare Veränderungen, z.B. hinsichtlich der Lehrmethoden, der Zugänglichkeit der Lehrenden oder der Abstimmung von Lehrangeboten im Fachbereich.

Diese Diskrepanz erklärt möglicherweise, wieso die Methode studentischer Veranstaltungskritik oft punktuell und einmalig geblieben ist. Offenbar sind viele Lehrende zwar neugierig, wie sie in den Augen der Studierenden gesehen werden, aber sie fassen studentische Beurteilungen nicht als Instrument permanenter Qualitätsmessung und Lehrverbesserung auf. Bei den Studierenden macht sich eine gewisse „Ermüdungstendenz“ breit, da sie ihre Erwartungshaltung auf eine Lehrverbesserung, die auch eine Änderung der Lehrmethoden oder gar der Inhalte umfassen würde, nur selten befriedigt sehen. Diese Feststellungen decken sich mit einer Befragung aller Dekane an nordrhein-westfälischen Hochschulen (MINISTERIUM FÜR WISSENSCHAFT UND FORSCHUNG 1998). Die Dekane geben an, daß sie studentische Befragungen vor allem als Kommunikationsmittel zwischen Studierenden und Lehrenden in einzelnen Veranstaltungen sehen und daran in der Regel keine weiteren Folgerungen geknüpft werden.

An dieser Zurückhaltung wird deutlich, daß sich eine Reihe von Rahmenbedingungen ändern muß, damit Evaluationsdaten tatsächlich zur Verbesserung der Lehre genutzt werden. Hauptansatzpunkt muß dabei die Stellung der Lehre an den Hochschulen sein (vgl. auch HELMKE/KRAPP 1999). Sie läßt sich

zu einem gewissen Teil durch Steuerungsmechanismen beeinflussen. Auch die Hochschulrektorenkonferenz rechnet damit, daß sich Evaluationsmaßnahmen und ein „einheitliches System der Qualitätssicherung“ (HOCHSCHULREKTOREN-KONFERENZ 1998, S. 13) als akzeptierte Verfahren an den Hochschulen durchsetzen werden. Darüber hinaus wird erwartet, daß Evaluationsdaten auch bei der langfristigen Strategieplanung, der Mittelverteilung, der Reform von Leitungsstrukturen, der Organisationsentwicklung und der Akkreditierung von Studienprogrammen und Institutionen verwendet werden. Dies alles sind bislang aber noch ungesicherte Zukunftsperspektiven hinsichtlich des Stellenwertes der Lehrevaluation an den Hochschulen.

Die Evaluation der Lehre hat sich zumindest als eine wichtige Voraussetzung herausgestellt, um das vernachlässigte Thema der Lehrqualität an den Hochschulen zu etablieren. Das ist über Rankings und andere Verfahren angeregt und vermittelt worden. Eine Reihe von Einwänden gegen die Verwendung studentischer Stellungnahmen können aufgrund vorhandener Empirie als unbegründet gelten. Allerdings bedarf es gut entwickelter Instrumentarien und angemessener Auswertungen. Um die Lehrevaluation nicht als Sackgasse erscheinen zu lassen, die letztlich nur wenig zur Verbesserung der Lehre beiträgt, muß sie in Ziel- und Kontrollsysteme eingebaut werden, wie sie im Rahmen von „Peer-Reviews“ sich zu entwickeln beginnen. Sie bedarf außerdem der unterstützenden Maßnahmen und Ressourcen, seien es institutionalisierte Evaluationsagenturen, Anreizsysteme oder Weiterbildungsangebote für Lehrende, soll ihre Wirkung nicht nur punktuell, sondern nachhaltig sein.

Literatur

- ALEAMONI, L.M.: Typical faculty concerns about student evaluation of teaching. In: *New Direction for Teaching and Learning* 31 (1987), S. 25–31.
- AMELANG, M./RINDERMANN, H.: Das Heidelberger Inventar zur Lehrveranstaltungsevaluation (HILVE). Heidelberg 1994.
- BARGEL, T.: Studienqualität und Hochschulentwicklung. (Bildung-Wissenschaft aktuell 11/93) Bonn 1993.
- BARGEL, T.: Studentische Erwartungen an die Lehr- und Studienqualität – Thesen und Erläuterungen. In: TECHNISCHE UNIVERSITÄT DRESDEN (Hrsg.): *Evaluation der Lehre*. Dresden 1996, S. 97–104.
- BARGEL, T./MULTRUS, F./RAMM, M.: Studium und Studierende in den 90er Jahren. Entwicklung an Universitäten und Fachhochschulen in den alten und neuen Bundesländern. Bonn 1996.
- BEATTY, M.J./ZAHN, C.J.: Are student ratings of communication instructors due to „easy“ grading practices? An analysis of teacher credibility and student-reported performance levels. In: *Communication Education* 39 (1978), S. 275–282.
- BERENDT, B./STARY, J. (Hrsg.): *Evaluation zur Verbesserung der Qualität der Lehre und weitere Maßnahmen*. Weinheim 1993.
- BÜLOW-SCHRAMM, M.: Wer hat Angst vor den Evaluatoren? Der Umgang mit Akzeptanzproblemen von Evaluationsverfahren. In: *Handbuch Hochschullehre Highlights, Evaluation der Lehre*. Band 1: Ziele – Akzeptanz – Methoden. Stuttgart 1995, D 1.6, S. 1–16.
- CENTRA, J.A.: *Determining faculty effectiveness*. San Francisco 1980.
- CHRISTIAN, D.E.: *Social Indicators. The OECD Experience*. Paris 1974.
- COMELLY, G.: Studentische Veranstaltungskritik. Mit wachen Augen trotzdem mitmachen! In: *Die Neue Hochschule* 4 (1991), S. 12–14.
- DANIEL, H.-D./THOMA, M./BANDILLA, W.: Das Modellprojekt „Evaluation der Lehre“ an der Universität Mannheim. In: MOHLER, P. (Hrsg.): *Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung*. Münster 1994, S. 81–90.
- DANIELSEN, A.L./WHITE, R.A.: Some evidence on the variables associated with student evaluation of teachers. In: *The Journal of Economic Education* 7 (1976), S. 117–119.

- DIEHL, J.M.: Fragebogen zur studentischen Evaluation von Hochschulveranstaltungen. Manual zum VBVR und VBREF. (Fachbereich 06 Psychologie (Abteilung Methodik), Universität Gießen) Gießen 1994.
- EITEL, F.: Lernforschung als Voraussetzung für die Unterrichtsorganisation. In: BICHLER, K.H./MATTAUCH, W./WECHSEL, H.W. (Hrsg.): Innovationen und Trends des Medizinstudiums im klinischen Teil. Band II. Frankfurt a.M. 1995, S. 72–96.
- ELMORE, P. B./POHLMANN, J. T.: Effect of teacher, student, and class characteristics on the evaluation of college instructors. In: *Journal of Educational Psychology* 70 (1978), S. 187–192.
- ESSER, H.: No Shows, Karteileichen, Schleifendrehen. In: *Deutsche Universitätszeitung* 18 (1995), S. 22–25.
- FOCUS: Deutschlands größter Uni-Test. Deutschlands Top-Universitäten, In: *Focus* 16 (1997), S. 138–158.
- GIESEN, H./JANSEN, R.: Universitäten aus der Sicht ihrer Studenten. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 15 (1983); S. 222–233.
- GRÜHN, D./GATTWINKEL, H. (Hrsg.): Evaluation von Lehrveranstaltungen. Überfrachtung eines sinnvollen Instruments? (Informationen aus Lehre und Forschung. 2). Berlin 1992.
- HAGE, N. EL: Lehrevaluation und studentische Veranstaltungskritik. Projekte, Instrumente und Grundlagen. Bonn 1996a.
- HAGE, N. EL: Studentische Urteile zur Lehr- und Studienqualität. Analysen zu ihrer Validität anhand der Indikatoren und Skalen im Studierendensurvey. (Hefte zur Bildungs- und Hochschulforschung. Bd. 19), Konstanz 1996b.
- HECKHAUSEN, H.: Die Pädagogische Psychologie vor neuen Herausforderungen. In: B. WEIDEMANN/A. KRAPP (Hrsg.): *Pädagogische Psychologie*. Schwarzberg 1986, S. 786–788.
- HELMKE, A.: Studentische Evaluation der Lehre – Sackgassen und Perspektiven. In: *Zeitschrift für Pädagogische Psychologie* 10 (1996), H. 3/4, S. 181–186.
- HELMKE, A./KRAPP, A.: Lehren und Lernen in der Hochschule. Einführung in den Thementeil. In: *Zeitschrift für Pädagogik* 45 (1999), H. 1, S.19–24.
- HOCHSCHULREKTORENKONFERENZ: Evaluation. Sachstandsbericht zur Qualitätsbewertung und Qualitätsentwicklung in deutschen Hochschulen. Projekt Qualitätssicherung. (Dokumente & Informationen Bd. 1). Bonn 1998.
- HOFMANN, J. M.: Studienmotivation und Veranstaltungsbeurteilung. In: *Psychologie, Erziehung, Unterricht* 35 (1988), S. 119–126.
- HOLTkamp, R./SCHNITZER, K.: Evaluation des Lehrens und Lernens. Ansätze, Methoden, Instrumente. Evaluationspraxis in den USA, Großbritannien und den Niederlanden. Dokumentation der HIS-Tagung am 20./21. Februar 1992. In: *HIS Hochschulplanung*, Bd. 92. Hannover 1992.
- HORNBOSTEL, S./DANIEL, H.-D.: Das Spiegel-Ranking: Mediensensation oder ein Beitrag zur hochschulvergleichenden Lehrevaluation? In: MOHLER, P. (Hrsg.): *Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung*. Münster 1994, S. 29–44.
- KROMREY, H.: Studentische Vorlesungskritik. Empirische Daten und Konsequenzen für die Lehre, In: *Soziologie* 1 (1993), S. 39–56.
- KROMREY, H.: Evaluation der Lehre durch Umfrageforschung? Methodische Fallstricke bei der Messung von Lehrqualität durch Befragung von Vorlesungsteilnehmern. In MOHLER, P. (Hrsg.): *Universität und Lehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung*. Münster 1994, S. 91–114.
- KROMREY, H.: Studentische Befragung zu Lehre und Studium. Von der Lehrevaluation zur Qualitätsentwicklung. In: TECHNISCHE UNIVERSITÄT DRESDEN (Hrsg.) *Evaluation der Lehre*. Dresden 1996, S. 67 – 87.
- LOCUMER PROTOKOLLE: Lehre an Hochschulen. Über Kriterien und Instrumente zu ihrer Evaluation und Förderung. Rehbürg-Loccum 1992.
- MANAGER MAGAZIN: Die besten Unis für Juristen. Abschied von gestern: 24 (1994), S. 138–167.
- MARSH, W.: Validity of students' evaluation of college teaching. A multitrait-multimethod analysis. In: *Journal of Educational Psychology* 74 (1982), S. 264–279.
- MARSH, W./WARE, J.E.: Effects of expressiveness, content coverage, and incentive on multidimensional student rating scales: New interpretation of the Dr. Fox effect. In: *Journal of Educational Psychology* 74 (1982), S. 126–134.
- MINISTERIUM FÜR WISSENSCHAFT UND FORSCHUNG DES LANDES NORDRHEIN-WESTFALEN (Hrsg.): *Reformen und Ressourcen. Zwischenbilanz und Perspektiven des Aktionsprogramms „Qualität der Lehre“*. Düsseldorf 1998.
- MURRAY, H.G.: Low-interference classroom teaching behaviors and student ratings of college teaching effectiveness. In: *Journal of Medical Psychology* 75 (1983), S. 138–149.
- NAFTULIN, D.H./WARE, J.E./DONNELLY, F.A.: The Dr. Fox lecture: A paradigm of educational seduction. In: *Journal of Medical Psychology* 48 (1973), S. 399–404.

- RCDS (Hrsg.): Mehr Ehre für die Lehre. Das neue Hochschulranking. Erlangen 1992.
- REISSERT, R.: Interne Selbstevaluation und externe Begutachtung durch Peers – Bericht über das Pilotprojekt Hamburg und Kiel. In: TECHNISCHE UNIVERSITÄT DRESDEN (Hrsg.): Evaluation der Lehre. Dresden 1996, S. 105–113.
- REISSERT, R./CARSTENSEN, D.: Praxis der internen und externen Evaluation. Handbuch zum Verfahren. In: HIS (Hrsg.): HIS-Kursinformationen. Hannover 1998.
- REUKE, H.: Flächendeckende und systematische Evaluation von Lehre und Studium in Niedersachsen. In: Beiträge zur Hochschulforschung 2 (1996), S. 81–100.
- RICHTER, R. (Hrsg.): Qualitätssorge in der Lehre. Leitfaden für die studentische Evaluation. Neuwied 1994.
- RICHTER, R.: Erfahrungen mit der internen und externen Qualitätsevaluation der Lehre an den niederländischen Universitäten nach Beendigung des ersten Evaluationszyklus 1993. In: TECHNISCHE UNIVERSITÄT DRESDEN (Hrsg.) Evaluation der Lehre. Dresden 1996, S. 115–126.
- RINDERMANN, H.: Untersuchungen zur Brauchbarkeit studentischer Lehrevaluationen. Landau: 1996.
- SCHOTT, E.: Zur empirischen und theoretischen Grundlegung eines Bewertungsinstrumentes für Vorlesungen. Blickpunkt Hochschuldidaktik, Bd. 28, 1973.
- SÖLLNER, A.: Von der Vermessung der Quantität zur Erprobung neuer Qualität. In: TECHNISCHE UNIVERSITÄT DRESDEN (Hrsg.): Evaluation der Lehre. Dresden 1996, S. 11–16.
- SPIEGEL: Die neuen Unis sind die Besten. Heft 50 (1989), S. 70–87.
- SPIEGEL (Hrsg.): Studieren heute – Welche Uni ist die beste? (SPIEGEL-SPEZIAL Nr.1) Hamburg 1990.
- SPIEGEL (Hrsg.): Welche Uni ist die beste? (SPIEGEL-SPEZIAL Nr.31) Hamburg 1993.
- SPIEGEL: Deutsche Hochschulen im Vergleich. Heft 15 (1999), S. 58–107.
- STANGL, C./MEISNER, H.-K./SELBMAN, H.-K./MOELLER, H.: Evaluierung der Qualität der Lehre aus der Sicht der Studierenden: eine interventionelle Befragungsstudie in Tübingen. In: BICHLER, K.-H./MATTACH, W./WECHSEL, H.W. (Hrsg.): Innovationen und Trends des Medizinstudiums im klinischen Teil (Band II). Frankfurt a.M. 1995, S. 97–105.
- TIMES: The Good University Guide. 23.3.1999.
- WARE, J.E./WILLIAMS, R.G.: The Dr. Fox effect: A study of lecturer effectiveness and rating of instruction. In: Journal of Medical Education 50 (1975), S. 149–156.
- PETER, L./WAWRZINEK, A.: Dialogische Evaluation. Ein studentisches Evaluationsverfahren. In: Handbuch Hochschullehre Highlights, Evaluation der Lehre. Band 1: Ziele – Akzeptanz – Methoden. Stuttgart 1995, D 3.1, S. 1–18.
- WEBLER, W.-D.: Evaluation im Kontext der Organisationsentwicklung. Erfahrungen mit einem Modell für Lehrberichte. In: Beiträge zur Hochschulforschung 3 (1995), S. 293–326.
- WEBLER, W.-D./DOMEYER, V./SCHIEBEL, B.: Lehrberichte: empirische Grundlagen, Indikatorauswahl und Empfehlungen zur Darstellung der Situation der Lehre in Lehrberichten. Bad Honnef 1993.
- WEINERT, F.E.: Lernforschung als eine zentrale Aufgabe der Pädagogischen Psychologie. In: WEIDEMANN, B./KRAPP, A. (Hrsg.): Pädagogische Psychologie. Schwarzberg 1986, S. 783–786.
- WINTELER, A./KRAPP, A.: Programme zur Förderung der Qualität der Lehre an Hochschulen. In: Zeitschrift für Pädagogik 45 (1999), H. 1, S.45–60.
- WISSENSCHAFTSRAT: Empfehlung zur Stärkung der Lehre in den Hochschulen durch Evaluation. Berlin 1996.
- WOLTER, A.: Evaluierung der Hochschullehre – Mehr Staatskontrolle oder mehr Hochschulautonomie? In: GNAHS, D./KREKEL, E.M./ WOLTER, A. (Hrsg.): Qualitätsmanagement im Bildungswesen. St. Marienthal 1996, S. 73–106.
- ZEITSCHRIFT FÜR PÄDAGOGISCHE PSYCHOLOGIE: Themenheft Lehr-Evaluation 10. 1996.

Anschrift der Autoren

Tino Bargel, Dipl. Soz., Sonneneck 1, 78465 Konstanz

Natalija El Hage, Landgrafenstr. 39, 60486 Frankfurt a.M.