Evaluation der Lehre: Praxiserfahrungen und Methodenhinweise

Wolff-Dietrich Webler

Seit der öffentlichkeitswirksamen Spiegelumfrage gibt es zahlreiche Initiativen an Hochschulen, in denen versucht wird, die Lehrveranstaltungen in möglichst vielen Studiengängen zu evaluieren. Dies geschieht mit sehr verschiedenen Erwartungshaltungen.

Auch die dabei angewendeten Evaluationsmethoden sind auf sehr unterschiedlichem methodischen Anspruchsniveau angesiedelt. Die wichtigste Informationsquelle sind die teilnehmenden Studierenden, deren Urteilsfähigkeit inzwischen aufgrund der Ergebnisse der empirischen Forschung nicht mehr angezweifelt wird.

Zu dieser Methodenvielfalt gibt der vorliegende Beitrag einen Überblick. Neben Hinweisen auf bestehende Initiativen an Hochschulen und bereits vorhandene Erfahrungen gibt der Autor schließlich noch Einblick in einige Strategien und Instrumente der Lehrevaluation und Erstellung von Lehrberichten am Fallbeispiel der Uni Bielefeld.

1 Kontext der Lehrevaluation

Evaluation der Lehre scheint das Thema der Stunde zu sein. Es bildet quasi die zweite Stufe der Debatte um Qualität der Lehre, die in den Bundesländern vor etwa drei Jahren begonnen hat. Zwischen 1966 und 1970/72 gab es in der Bundesrepublik Deutschland schon einmal eine Debatte um Veranstaltungskritik, die aber - von Studierenden ausgehend - vor allem wissenschaftstheoretische, wissenschaftskritische Fragen der Lehrinhalte im Vordergrund sah. Diese Debatte kam weitgehend zum Stillstand, bis am Ende der 80er Jahre die Frage der Qualität der Lehre neu gestellt wurde. In einer Gesellschaft, in der der sparsame Umgang mit öffentlichen Mitteln immer mehr mit Maßstäben des betrieblichen Rechnungswesens gemessen wird, stehen auch die Hochschulen unter dem öffentlichen Legitimationsdruck, die Erfüllung ihrer Aufgaben (hier: die Erzeugung gesellschaftlich notwendiger Qualifikationen) in Relation zum Verbrauch öffentlicher Ressourcen zu setzen. Und hier haben die Hochschulen - berechtigt oder unberechtigt - sich einen üblen Ruf zugezogen: Die Finanzministerien und Wissenschaftsausschüsse der Landtage halten die Hochschulen inzwischen in weiten Teilen für einen zu rationalen Prioritätenentscheidungen unfähigen, Kostengesichtspunkten gegenüber verständnislosen, sich immer neue Privilegierungen verschaffenden, ineffizienten "Chaoshaufen", ein undurchsichtiges Millionengrab. Dieser Eindruck wird in Fragen des Studiums insbesondere an den hohen Abbrecher- und Fachwechslerquoten festgemacht. Als Zustand der Lehre sehen sie:

1. einen ineffizienten Personal-/Mittelleinsatz,
2. unnötige/ungewünschte Zeitverluste im Studium,
3. Fehlinvestitionen von Ausbildungszeit und Ausbildungsressourcen durch unnötig hohe Abbrecher-/Fachwechslerquoten,
4. einen zu geringen Lehr-/Lernerfolg im Sinne von Studienerfolg.

Schon macht die Forderung nach "erfolgsabhängiger Finanzierung der Hochschulen" in Wirtschaft und Landtagen die Runde. Auch wenn über Ursachen und Wirkungen trefflich gestritten werden kann, ist klar, daß die Hochschulen sich bezüglich ihrer Lehrleistungen öffentlichen Legitimationserwartungen gegenübersehen. Dabei geht es zunächst nicht um die Frage, ob Defizite bestehen oder nicht, sondern darum, daß Verfahren der Qualitätssicherung entwickelt und praktiziert werden, die als Sensor wirken. Dies ist keine spezifisch deutsche Entwicklung. Seit 1985 haben Frankreich, die Niederlande, Belgien und Großbritannien solche Systeme der Qualitätssicherung entwickelt; die skandinavischen Länder folgen zur Zeit. Dabei wirken sich Instrumente und Verfahren der Qualitätssicherung in der Lehre im internationalen Vergleich als wichtige Elemente der institutionellen Selbstregulie-

rung im Rahmen einer effektiven Wahrnehmung der Hochschulautonomie aus. - Dies ist der politische Rahmen, in dem die Forderung nach einer Evaluation der Lehre gesehen werden muß.

Daneben existiert der im engeren Sinne hochschuldidaktische Rahmen, der von den beteiligten Individuen, den Lehrenden und Lernenden, ausgeht. Sich selbst über den Erfolg eigenen, zielgerichteten Tuns zu vergewissern, ist in vielen Lebensbereichen eine Selbstverständlichkeit. Im Rahmen akademischer Lehre gilt dies nur höchst eingeschränkt. Zwar können alle Formen der Leistungsmessung (von ihrer Bedeutung im Rahmen des Zertifizierungs- und Berechtigungswesens abgesehen) auch als eine Form des feed-back zum Lehrerfolg der Lehrenden interpretiert werden. Die Informationen bleiben allerdings in der Regel auf kognitive Leistungen beschränkt. Klausuren und alle anderen Formen von Prüfungen stellen allein auf die Feststellung des Lernerfolgs der Studierenden ab und lassen die Anteile der Lehre an Erfolg oder Mißerfolg systematisch unberücksichtigt. Ein anderer Aspekt kommt hinzu: Die Studierenden diskutieren die andere Seite des Lernprozesses, die Lehrangebote und ihre Qualität durchaus. In informellen Gesprächen und durch die studentische Studienberatung der Fachschaften und Hochschulgruppen werden Einschätzungen zu Veranstaltungen und Personen weitergegeben. Der Ruf als Lehrende(r) ist Ergebnis einer nichtöffentlichen, unsystematischen und intransparenten Sammlung von Impressionen und Erfahrungen, in der die Lehrperson das Objekt der Aufmerksamkeit, aber nicht handelndes Subjekt ist - einer Sammlung, deren Ergebnisse sich im Zulauf oder Wegbleiben der Studierenden, aber kaum je als Ergebnis einer Kommunikation über Lehren und Lernen niederschlagen. Derartige Reaktionen lassen die Lehrenden nicht unberührt, geben ihnen aber kaum Anhaltspunkte für die Ursachen, Fehlinterpretationen sind oft die Folge. Erst das Gespräch gäbe beiden Seiten die Möglichkeit der Information und Erklärung. Die stattdessen oft herrschende Sprachlosigkeit auf diesem Gebiet tut der nach wie vor gewollten Gemeinschaft der Lehrenden und Lernenden Abbruch und mindert Lehr- und Lernerfolg. Erwartungen der Lehrenden und Lernenden differieren oft erheblich und bleiben mangels wechselseitiger Kenntnis vielfach unerfüllt. Selbstbilder der Lehrenden über ihren Lehrerfolg entsprechen mangels (eigener) Erfolgskontrolle zu häufig nicht der Realität.

Voraussetzung für jede weitere Vertiefung der Kenntnisse ist allerdings das Bedürfnis der Lehrenden, mehr über die Ausgangsbedingungen und Wirkungen ihrer Tätigkeit zu erfahren. Dann besteht die Bereitschaft, innerhalb und außerhalb der Veranstaltung Zeit für die Evaluation vorzusehen und sie als nützlich für das eigene Lernen wie das der Teilnehmer anzuerkennen. Dieser Nutzen wird besonders gesteigert, wenn die Rückmeldung und Auswertung

festen Bestandteil der weiteren Ausgestaltung des Veranstaltungs- und Arbeitsablaufes ist; sie gewinnt dann auch für die Teilnehmer größere Bedeutung. Sensibler zu werden für den Ablauf von Lehr- und Lernprozessen, individuelle und Gruppenarbeitsprozesse beobachten, durchdenken und steuern zu können, ist nicht nur in sozialwissenschaftlichen Studiengängen im weitesten Sinne heute schon Teil einer Berufsqualifikation, sondern auch darüber hinausgehend ein Lerngewinn. Gemeinsam gewonnene Erkenntnisse dieser Art können zur Grundlage einer vertrauensvollen, gemeinsamen, auf gegenseitige Anerkennung beruhenden Lösung auftauchender Probleme des Lehrens und Lernens werden und einen eventuellen Zerfall von Lehrveranstaltungen in Konsumenten- und Lieferantenrollen in sozialer Distanz verhindern, wie er sich zeitweise auszubreiten scheint.

Soll nicht nur die Lernleistung beleuchtet, sondern auch die Lehraufgabe der Hochschule in ihrer Umsetzung transparent gemacht, soll die Adäquanz der gewählten Didaktik für den Weg zwischen (selbst-)gesetzten Lehrzielen und der Ausgangssituation der StudienanfängerInnen bzw. des jeweiligen Semesterbeginns sowie der Erfolg der Lehrenden für eine Selbstkontrolle, Revision und Selbststeuerung festgestellt werden, so müssen spezifische Verfahren und Instrumente zur Rückmeldung des Lehrerfolgs eingesetzt werden. Diese Verfahren, in denen empirisch Informationen über Bedingungen, Systeme, Konzepte, Verläufe und Wirkungen von Lehr- und Lernprozessen gewonnen werden, verbunden mit der expliziten Aus- und Bewertung der Informationen, werden mit dem Begriff der Evaluation zusammengefaßt.

2 Ziele der Lehrevaluation

Im Vordergrund der Rückmeldung des Lehr- und Lernerfolgs steht die interne Optimierung der Veranstaltung und auch der Fortbildungszweck für die Lehrenden, die dann eigene Schlüsse zur Verbesserung ihrer didaktischen Kompetenz aus den Ergebnissen der Evaluation ziehen können. Daher findet eine solche Evaluation etwa in der Mitte des Semesters statt.

Als zweites erfüllt die Evaluation das Ziel der Förderung der Kommunikation mit Studierenden, also zwischen Lehrenden und Studierenden sowie unter Studierenden, indem beide Seiten das erste Mal über Material verfügen, das die Analyse des Lehr- und Lernprozesses ermöglicht.

Drittens geht es um die Förderung der Lernmotivation und Lernfähigkeit der Studierenden. Hierbei ist der Effekt nicht zu unterschätzen, daß Studierende durch die Evaluation auf wichtige Dimensionen des Lehrprozesses aufmerksam gemacht werden. Fragebögen sind immer gleichzeitig For-

schungs- und auch didaktische Instrumente. Die Befragten werden zunächst einmal auf Fragenswertes aufmerksam gemacht, bevor sie Antworten formulieren.

Das vierte Ziel ist die Studienberatung, die Auswahl der Kurse und Lehrenden. Die Studierenden erhalten Anhaltspunkte über Stärken und Schwächen und können bei Wahlmöglichkeiten ihre Entscheidungen entsprechend einrichten.

Der fünfte Zweck ist die Verbesserung der Effizienz des Studiengangs. Die Evaluation dient auch als Grundlage zur Veränderung des Lehrangebotes. Wenn sich Mängel, die vielleicht nicht in der Lehrperson, sondern in der Struktur dieser Veranstaltung liegen, herausstellen, dann kann das Evaluationsergebnis als Grundlage zur Veränderung des Angebots herangezogen werden.

3 Methodische Alternativen

Die vorliegenden Evaluationsmethoden sind sehr variabel und je nach Verwendungszusammenhang auf sehr unterschiedlichem methodischen Anspruchsniveau angesiedelt. Zwar gibt es die Möglichkeit für Lehrpersonen, sich eine Beurteilung der eigenen Lehre durch hospitierende Kollegen, durch team-teaching, beobachtende Hilfskräfte oder die Selbstausswertung eines Videobandes der eigenen Veranstaltung zu verschaffen, aber die wichtigste Informationsquelle sind die teilnehmenden Studierenden. Gegen deren Urteilsfähigkeit werden zwar eine Reihe von Zweifeln vorgebracht (insbesondere im Zusammenhang mit Schätzverfahren; vgl. Winteler, A. 1978): Studierende beurteilten in Abhängigkeit von erhaltenen Zensuren; Beurteilungsgegenstand sei eher die Popularität oder Show eines Dozenten als der Unterricht; aus mangelnder Erfahrung seien sie wohl fähig, über die Lehrbefähigung zu urteilen, der Zusammenhang von Lehre und Forschung werde durch Studierende jedoch nicht erfaßt; die Zustimmung der Mehrheit sei kein Gütesiegel für gute Lehre; es fände eine oberflächliche, nicht vorurteilsfreie Urteilsbildung statt; studentische Urteile seien nicht zuverlässig und gültig; Schätzskalen sorgten für Unruhe und Spannung zwischen Lehrenden und Lernenden und zwischen Lehrenden; die Entwicklung von Schätzskalen sei zu zeit- und geldaufwendig.

Die Ergebnisse der empirischen Forschung haben diese Bedenken nicht bestätigt (s. im einzelnen schon Winteler, A. 1978, S. 68 ff.). Als gesichert kann demgegenüber gelten, daß die didaktische Gestaltung der Lehrveranstaltung ein wichtiger Faktor für ihre Effektivität ist (und somit die Gestaltung optimiert werden sollte). Auch der verbessernde Einfluß von Evaluationsver-

fahren (i.e.S.: Schätzverfahren) auf die Lehrwirksamkeit ist durch zahlreiche Studien belegt (ders., S. 83).

Es werden Evaluationsverfahren für Forschungszwecke (z.B. Evaluation der Bewährung eines neuen Studiengangs, detaillierte Studien zur Lehr-/Lernforschung u.ä.) und Alltagsverwendungen der Evaluation durch die direkt Betroffenen zur unmittelbaren Rückmeldung des Lehr-/Lernerfolgs und zur Steuerung der weiteren Prozesse im Veranstaltungsverlauf unterschieden. In diesem zweiten Fall muß das Procedere einfach genug sein, um es im Alltag des Semesters ohne unzumutbaren Aufwand und in der Regel ohne fremde Hilfe anwenden zu können. Es muß in der Regel auch ohne spezifische Kenntnisse der empirischen Sozialforschung in allen Fachrichtungen einsetzbar sein.

Wie der Verfasser an anderer Stelle ausgeführt hat (Webler 1983), muß Evaluation im Regelfall die Ermittlung der Wechselwirkungen zwischen den situativen Bedingungen sowie den in den arrangierten Lernsituationen interagierenden Personen (Lehrenden und Lernenden) einschließen. Weil die didaktischen Interaktionen zielbestimmt sind, muß die Evaluation allerdings vor allem diejenigen Bedingungen/Eigenschaften der Lernsituationen ermitteln, die geeignet sind, die aufgestellten Ziele zu erreichen sowie diejenigen, deren Auswirkungen auf Verhalten und Kognitionen der Beteiligten im Sinne der Zielsetzungen interpretierbar sind. Eine solche Evaluation geht nach dem Muster kommunikativer Sozialforschung vor, die von der Strukturierung von Situationen und Prozessen sowie der kommunikativen Aufarbeitung von Erfahrungen bestimmt ist. Hierbei kommt es durchaus vor, daß die Sicherung einer empirischen Grundlage und der systematischen Auswertung der Erfahrungen vernachlässigt werden.

In einer streng empirisch vorgehenden Evaluation werden die Beziehungen zwischen aufgestellten Lernzielen, gewählten Lernsituationen und deren Ablauf als "didaktische Hypothesen" aufgefaßt. Die Überprüfung der didaktischen Hypothesen geschieht durch die systematische Gewinnung von Informationen über Bedingungen, über Konzepte von Lernprozessen und ihre tatsächlichen Verläufe und über Ergebnisse und Wirkungen von Lehr- und Lernprozessen bei Lehrenden, Lernenden oder sonstigen Betroffenen (Kollegen, institutionellen Trägern etc.). Das Ziel der Überprüfung der didaktischen Hypothesen ist die Formulierung von Bedingungen, unter denen die aufgestellten Ziele mit mehr oder weniger großer Wahrscheinlichkeit zu erreichen sind. Im Sinne einer rationalen Begründung von didaktischen Entscheidungen soll die Evaluation zeigen, daß bei Übernahme der Zielsetzungen zu deren Realisierung regelmäßig bestimmte Entscheidungen zu treffen und bestimmte Lernsituationen vorzusehen sind.

Die Entwicklung von Programmen sowohl zur Datengewinnung als auch zur Bewertung und Interpretation dieser Daten kann nun verschiedene Funktio-

nen erfüllen, die jedoch in allen Fällen handlungs- und damit entscheidungsorientierende Wirkung haben. Je nach dem Erkenntnisinteresse und dem Verwendungszweck sind die Evaluationsprogramme unterschiedlich zu akzentuieren; dabei sind verschiedene Funktionen möglich:

- Rekonstruktion und Begründung der Entwicklung von Curricula zwecks Einführung,
- Gewinnung von Material für die Curriculumrevision,
- Information zur Nachsteuerung des eigenen/fremden Lehr-/Lernprozesses,
- Steigerung der Lernmotivation,
- Revision einzelner Elemente (Verhalten, Inhalte, soziales Arrangement),
- Legitimation z.B. von curricularen Merkmalen und didaktischen Prinzipien (integrierte Studiengänge, Projektstudium etc.), aber auch (u.U. ungewollt):
- Erstellung von Rangreihen und Selektion (z.B. von Studenten) sowie
- Disziplinierung (der Teilnehmer und Träger) in Verbindung mit drohenden Sanktionen (z.B. eines Modellversuches).

Aus dem Spektrum vorhandener Möglichkeiten der Lehrevaluation können für Alltagszwecke der Lehre fünf verschiedene Verfahren vorgeschlagen werden, die auch kombinierbar sind (in der Reihenfolge des Anspruchsniveaus):

1. "Manöverkritik" pro Sitzung (letzte 5 - 20 Minuten),
2. Diskussion des Veranstaltungsverlaufs am Semesterende (1 Sitzung),
3. Verteilung von Beobachterrollen mit Beobachtungsbögen pro Sitzung,
4. Vergleich des (relativ detailliert fixierten) Veranstaltungsplanes mit dem Ablauf anhand von Sitzungsverlaufsprotokollen,
5. Fragebogen-Erhebungen zu folgenden, kombinierbaren Erkenntnisinteressen:
 - 5.1 zu den Erwartungen der Studierenden,
 - 5.2 zu den Lernvoraussetzungen (insbesondere Vorkenntnissen) der Studierenden,
 - 5.3 zu Interaktionsprozessen/zu den gewählten Lernsituationen,
 - 5.4 zum Lernverhalten der Studierenden,
 - 5.5 zum Lernverhalten der Lehrpersonen,
 - 5.6 zu schriftlichen Materialien (Auswahl von Texten, statistischen Materialien, Abbildungen, Handzetteln) und sonstigen Medien,

- 5.7 zu den erreichten Lehrzielen (Stoffverständnis, Wissen, praktische Anwendungs- und Übertragungsfähigkeit, Teamfähigkeit, kommunikative Kompetenz etc.).

Die sogenannte **Manöverkritik** dient (als Meta-Kommunikation über Lehren und Lernen) einer unmittelbaren Aussprache zwischen Lehrenden und Lernenden als Rückmeldung über alle Aspekte des Veranstaltungsverlaufs. Sie ermöglicht die sofortige Aufklärung von Mißverständnissen, fördert eine engere Abstimmung zwischen Lehrintention und Lernabläufen (Tempo des Vorgehens, gewünschte Vertiefung, Beispiele, spezifische Wiederholungen) und erhöht die beiderseitige Zufriedenheit als zwei wichtige Faktoren des Lehr-/Lernerfolgs. In Veranstaltungen mit großer Teilnehmerzahl können gegenseitige Interviews von Nachbarn (**Paar-Interviews**) eine Bündelung der "Manöverkritik" vorbereiten oder der Einübung der Reflexion über Wünsche an die Lehre, Veranstaltungspläne und -abläufe dienen, in größeren (Seminar-)Veranstaltungen evtl. auch durch den Einsatz der **Gruppendiskussionsmethode**.

Die **Diskussion des Veranstaltungsverlaufs am Semesterende** (anhand einer Rekonstruktion oder des Textes im kommentierten Vorlesungsverzeichnis u.ä.) verläuft geraffter, abstrakter als im ersten Beispiel und kann als Bilanzierung keine laufende Kurskorrektur des Ablaufs mehr bieten, aber wichtige Hinweise zur künftigen Gestaltung von Lehrveranstaltungen geben. Es sollte nicht nur Negativkritik geübt werden, sondern positive Seiten herausgearbeitet und im übrigen konstruktive Verbesserungsvorschläge vorgelegt werden.

Die Verteilung von **Beobachterrollen** mit schriftlicher Protokollierung oder eine Videoaufzeichnung ist zu empfehlen, wenn anhand der Kommunikationsstrukturen z. B. die Ursachen für das konstante Schweigen einzelner, schleppende Diskussion, extreme Lehrerzentrierung des Verlaufs oder das Erreichen kommunikativer Lehrziele der Veranstaltung erfaßt werden sollen. Alternativ können die Rollen einer Beobachtung durch Teilnehmer im Wechsel (ungeschult) oder durch Tutoren/Hilfskräfte kontinuierlich (dann etwas mehr vorbereitet) übernommen werden. Eine inzwischen häufiger geübte Variante ist das **Video-Feedback**, d.h. also das Aufnehmen einer Sitzung mit einer oder mehreren Kameras. In einer Nachbereitung wird das Videoband angesehen und interpretiert. Die Effekte sind äußerst wirksam, sowohl bei Lehrenden wie Lernenden, weil die Selbstbilder meist von der Realität abweichen.

Eine ebenfalls einfache Art der Soll-/Ist-Bilanzierung und Verlaufskontrolle besteht im regelmäßigen Vergleich der **Veranstaltungsplanung mit dem**

konkreten Verlauf. Dazu ist eine vorherige Formulierung der Veranstaltungsziele (Lehrziele) und der gedachten Etappen auf dem Weg dorthin (Gliederung der Veranstaltung und eingesetzte Medien, Lernsituationen usw.) sowie eine Protokollierung des tatsächlichen Verlaufs notwendig. Diese Vermutungen über den Zusammenhang zwischen Lehrzielen und geplantem Verlauf können als "didaktische Hypothesen" bezeichnet werden, die mit ihrer Formulierung einer Überprüfung zugänglich werden.

Die gegenüber offenen, unstrukturierten Diskussionen meist ergiebigere Form der Vergewisserung über Lehr- und Lernerfolge ist die **schriftliche Befragung**, die wegen der größeren Anzahl der Antwortenden und ihrer Anonymität ein zuverlässigeres Meinungsbild erzeugt. Entsprechend der Variabilität der Didaktik, d.h. der Kombination der Lehrziele und Lehrmethoden können Befragungen die unterschiedlichsten Erkenntnisinteressen verfolgen. Je nach Erkenntniszweck können Befragungen nur am Anfang des Semesters (z.B. zum Kennenlernen der Lernvoraussetzungen der Studierenden, des Kenntnisstandes - anonym - und zum Abstimmen des eigenen Vorgehens), am Anfang und am Ende (zur Feststellung des Lehrerfolgs als Vorher/Nachher-Effekte) oder im Verlauf in Frage kommen. Die verschiedenen Möglichkeiten können hier nur angedeutet werden. Das einfachste, aber erstaunlich wirksame Mittel ist das sogenannte **one-minute-paper**, besonders in Vorlesungen zu empfehlen. Die Studierenden werden am Ende der Vorlesung aufgefordert, auf einem Zettel die ihrer Meinung nach wesentlichen Punkte - oder (je nach Wunsch der Lehrenden) den unklarsten Punkt ("the muddiest point") der heutigen Vorlesung aufzuschreiben. Im Hinausgehen werden diese Zettel in bereitstehende Körbe geworfen. Der Lehrende sieht sich die Zettel an und kann daraus Schlüsse ziehen, wie weit das, was er sich als Schwerpunkt vorgestellt hatte, auch tatsächlich verstanden wurde. Dementsprechend kann er auch seine Vorlesung umstellen. Das Ergebnis kann auch Auslöser einer Diskussion über eine Veränderung der Vorlesung sein.

Ausdifferenzierte Fragebögen bieten demgegenüber breitere Informationen zu unterschiedlichen Aspekten des komplexen Lehr-/Lernprozesses.

4 ·Zur Konstruktion von Evaluationsfragebögen

Für die Verwendung von Fragebögen stehen prinzipiell drei Wege zur Auswahl:

1. Einen **relativ offenen Fragebogen** einzusetzen (wie ihn z.B. die Harvard University verwendet; "offen" deswegen, weil Fragen nicht durch Ankreuzen vorgegebener Antworten, sondern durch eigene Formulierungen der Befragten beantwortet werden).
2. Einen **Fragebogen mit geschlossenen Fragen** zu wählen (meistens als Batterie von Aussagen, zu denen auf einer 5er- oder 7er Skala das Ausmaß der Zustimmung bzw. Ablehnung erhoben wird).
3. Einen **Fragebogen nach dem "Cafeteria-System"** zu organisieren; bei solchen Fragebögen stellen die Lehrenden (zusammen mit Studierenden) aus einem Pool von einschlägigen, getesteten Frageformulierungen (ca. 400) dasjenige Menü auf dem Computer zusammen, das für die spezifische Veranstaltung einschlägig erscheint.

Die Entscheidung für eines der Muster (nicht einen konkreten Fragebogen, weil keines der Beispiele einfach übertragbar ist) hängt u.a. von folgenden Überlegungen ab:

Zu 1. Wenn sich die Befragung auf eine unbekannte (gerade erst festzustellende) Variabilität von Lehr- und Lernformen bezieht, ist das Instrument so offen zu gestalten, daß es die ganze Vielfalt der Veranstaltungsabläufe abbilden kann, ohne daß die Realität durch unzureichende Fragekonstruktionen bei der Antwort bereits verfälscht wird. Zu beachten ist, daß der Harvard-Fragebogen (bei Sprachstudien mit einer Variante) trotz seiner offenen Fragen für alle Lehrveranstaltungen (auch Massenveranstaltungen) ausgegeben und von einer Kommission der Harvard University ausgewertet wird. Bezeichnenderweise wird der Wiedergabe der Kritik zunächst die offizielle Veranstaltungsbeschreibung mit den Zielen der Lehrveranstaltung vorangestellt (unserem kommentierten Vorlesungsverzeichnis etwa entsprechend). Die Aufarbeitung der Fragebogen-Ergebnisse zu einer Art Lehrbericht von ca. 650 Seiten wird den Studienanfängern des folgenden Studienjahres vor der Immatrikulation zur Studienberatung zur Verfügung gestellt. Ich empfehle keinesfalls die einfache Übernahme so vieler offener Fragen.

Zu 2. Dieser Fragebogentypus ist der rationellste und auswertungsfreundlichste. Er setzt aber eine genaue Kenntnis des Untersuchungsfeldes voraus und ist im allgemeinen nicht ohne weiteres für alle Veranstaltungstypen gleichzeitig anwendbar, da dann umständliche Schleifen ("wenn nicht zutreffend, dann weiter mit Frage 10...") eingebaut werden müssen.

Zu 3. Das Cafeteria-System ist dann geeignet, wenn ein einziges Instrument mit geschlossenen Fragen für eine Vielzahl unterschiedlicher Lehr- und Lernformen eingesetzt werden soll. Der "Menü-Aufbau" erleichtert die Auswertung und minimiert die Gefahr falscher Fragenkonstruktionen, da er auch den der empirischen Sozialforschung fernstehenden Fächern getestete Fragen zur Verfügung stellen kann.

5 Kriterien für Evaluationsinstrumente in der Lehre

Da diese Instrumente nicht einfach nur Prozesse vollständig abbilden (wie eine Videokamera), und die Auswertung und Interpretation dann erst beginnt, sondern Evaluationsinstrumente selektiv nur bestimmte Variablen erfassen, stellt die Auswahl der Variablen eine Relevanzentscheidung dar, die auf Annahmen über die Bedeutung der betreffenden Variablen für die Rekonstruktion und Erklärung der Prozesse und ihres Lehr- und Lernerfolgs beruhen, also auf didaktischen Hypothesen. Bei der Konstruktion eines standardisierten Fragebogens, die sich auf eine unbekannte (gerade erst festzustellende) Variabilität von Lehr- und Lernformen bezieht, ist das Instrument breit genug zu gestalten, so daß die Realität nicht durch unzureichende Fragekonstruktion bereits verfälscht wird.

Im vorliegenden Fall soll Lehr- und Lernerfolg festgestellt werden, denn die Qualität der Lehre, die hier im Mittelpunkt steht, bemißt sich nach dem Ausmaß der Zielerreichung; es kann aber auch die Güte von Lehrprozessen erfaßt werden.

In jede Konstruktion eines Evaluationsfragebogens gehen regelmäßig Annahmen über Merkmale guter Lehre mit ein. Sie werden meistens nicht explizit formuliert, bestimmen aber die Auswahl der Variablen. Merkmale guter Lehre zeigen sich (als didaktische Entscheidung) in der Auswahl von Inhalten und den korrespondierenden Methodenentscheidungen. Da aber die eingesetzten Lehrmethoden in ihrer Eignung nur einschätzbar sind, wenn die Frage "geeignet wofür?" beantwortet werden kann, ist deren isolierte Erhebung wenig sinnvoll, und auch hier stellt sich die Frage nach den Lehr- und Lernzielen.

Evaluationsinstrumente müssen also

- a) in ihrer Auswahl von Variablen den Lehr- und Lernprozeß so erfassen können, daß sie als Basis für Aussagen über die Qualität der Lehre in der Relation von eingesetzten Lehr-/Lernmethoden zu konkreten Zielen taugen;

- b) operabel, d.h. in großer Zahl mit begrenztem Aufwand einsetzbar und auswertbar sein;
- c) offen genug sein, um alle existierenden Varianten von Lehre auch erfassen zu können; eine einseitige Evaluation der Lehre führt mittelfristig zu ihrer einseitigen Weiterentwicklung. Große Sorgfalt ist daher geboten.

Für die Entwicklung eines möglichst vielseitig einsetzbaren Evaluationsinstruments schlage ich ein "Zwiebelmodell" vor:

1. Im Kern eine konstante, überall anwendbare Fragenbatterie.
2. Darum herum einen Kranz fach- und lokalspezifischer Fragen, mit denen das jeweilige lokale Profil der Lehre (wichtig für den Lehrbericht) herausgearbeitet werden könnte. Ein Teil dieser Fragen kann ebenfalls vorab formuliert werden. Ein anderer sollte aber vor Ort hinzugefügt werden, und dafür sollten den Fachbereichen einige Musterbeispiele und Anleitungen als Material zur Verfügung gestellt werden.
3. Als äußere Schicht sollten wenige offene Fragen gestellt werden, wie sie etwa die Harvard University in ihrem Fragebogen verwendet. Zwar sind offene Fragen aufwendig in der Auswertung. Aber die Fachdiskussion ist sich einig, daß gerade zu Beginn eines Forschungsprozesses (wenn noch Lücken in der Feldkenntnis vermutet werden müssen, die zu lückenhaften oder falschen Fragestellungen geführt haben) offene Fragen unverzichtbar sind, um die ganze Realität zu erfassen. Das Antwortmaterial läßt dann oft zu (im Sinne eines Pretests), im zweiten Durchgang bestimmte Fragen zu revidieren, neue geschlossene Fragen aufzunehmen und damit die Anzahl offener Fragen zu reduzieren.

Die Konstruktion eines Befragungsinstruments, das den Ansprüchen an eine differenzierte Leistungsbeurteilung von Lehrenden und ihren Lehrveranstaltungen genügt, müßte von folgenden Prämissen ausgehen:

1. Lehrveranstaltungen sind komplexe, anspruchsvolle Prozesse, denen man mit simplen Frageinstrumenten nicht gerecht werden kann.
2. Evaluationsfragebögen dieser Art dienen in erster Linie der internen Optimierung der Lehrveranstaltung, sind also Kommunikationsmittel zwischen Studierenden und Lehrenden. Erst in zweiter Linie können aus den Ergebnissen - in aggregierter Form - auch weitergehende

Schlüsse zur Situation der Lehre in einem Studiengang gezogen werden.

3. Jede(r) Lehrende ist sicherlich enttäuscht, wenn der Fragebogen wegen zu simpler oder kurzer Anlage von vornherein keine Chance bietet, in den Antworten den eigenen wissenschaftlichen Anspruch und die Ziele, die die Lehrenden in ihre Veranstaltungen gelegt haben, angemessen abzubilden. Er ist dann auch keine gute Grundlage für ein Gespräch zwischen Lehrenden und Lernenden.
4. Sollen die Lehrenden den Ergebnissen nicht zu summarische Urteile entnehmen, sondern Informationen gewinnen, aus denen sie konkrete Schlußfolgerungen zur Verhaltensänderung bzw. zur Änderung von Lehrmaterial, Veranstaltungsaufbau o.ä. entnehmen können, müssen die Fragen solche konkreten Rückschlüsse auch zulassen.
5. Fragebögen sind immer gleichzeitig Forschungsinstrumente und didaktische Instrumente: Sie erfragen nicht nur Antworten, sondern machen mit ihren Fragen auch erst einmal auf das Fragenswerte aufmerksam. Diese Konsequenz ist unvermeidlich, kann aber auch offensiv genutzt werden. Üblicherweise als Einschätzung von Aussagen (Zu- oder Ablehnung) aufgebaut, kann der Fragebogen auch zur Mitteilung verwendet werden. Da den Studierenden, besonders in jungen Semestern, in vielen Bereichen einer Lehrveranstaltung von Kritikern derartiger Befragungen Urteilsfähigkeit abgesprochen wird, können relevante Kriterien guter Lehre aufgezählt und dann deren Erfüllung im konkreten Fall abgefragt werden. Eine so aufgebaute Befragung trägt wesentlich zur Entwicklung von Standards guter Lehre bei Studierenden - und bei Lehrenden - bei.
6. Ein Fragebogen, der diesen Anforderungen (1-5) genügen will, muß selbst die erforderliche Vielseitigkeit aufweisen. Die dafür angemessene Länge muß dann in Kauf genommen werden. Kürze ist kein Selbstzweck - wenn es selbstverständlich auch Obergrenzen der Bearbeitungszeit gibt.
7. Aus methodischen Gründen des gleichen Erhebungszeitpunktes und der gleichen Umstände (konstante Erhebungssituation) ist es notwendig, die Fragebögen in der Mitte einer Lehrveranstaltung (Sitzung) auszuteilen und an Ort und Stelle ausfüllen zu lassen, auch wenn dies die Veranstaltung verzögert. Der Zweck ist wichtig genug. Bei der Alternative, Verteilung am Sitzungsanfang oder Sitzungsende oder den Bogen nach Hause mitzugeben, treten zu viele Störfak-

toren, Fluktuation etc. auf. Wegen dieser Form der Ausfüllung ist die Länge oder Kürze des Fragebogens kein Argument zur Steigerung des Rücklaufs, wie sonst üblich.

8. Eine so gründliche Befragung ist für Lehrende und Fachbereiche nur in größeren Abständen notwendig. Eine Wiederholung reicht nach der Anfangsbefragung alle vier Jahre, wenn Lehrberichte alle zwei Jahre erscheinen sollen. Dazwischen können aus Gründen der Auswertungsökonomie auch wesentlich kürzere Befragungen ausreichen.
9. Die Fragebogen-Konstruktion sollte sich weitgehend auf die Feststellung der subjektiven Wirkungen des Lehrverhaltens und auf Beobachtungen der Befragten beschränken; nicht mehr, aber immerhin auch nicht weniger - handelt es sich doch schließlich um die Adressaten der Lehre.
10. Weitergehende Schlüsse sind oft von der Beobachtung nicht gedeckt, z.B. die Aufforderung zur Einschätzung der Feststellung: Der Veranstalter ist fachlich nicht kompetent. Zutreffend wäre jedoch eine Formulierung: Der/die Lehrende ist auf die Veranstaltung schlecht vorbereitet. Die Einschätzung von Wirkungen bzw. die Feststellung der Zustimmung oder Ablehnung zu einer Aussage auf einer Schätzskala ist methodisch angemessen, nicht jedoch die Beurteilung der Lehrperson als Benotung auf einer Skala von 1 - 5 oder 1 - 6. Letzteres hat auch emotional verheerende Wirkungen bei den Lehrenden und beeinträchtigt die konstruktive Form der Rückmeldung des Lehrerfolgs an die Lehrenden und ihre Kooperationsbereitschaft erheblich.

Für die Auswertung von Fragebögen ist die Ermittlung statistischer Mittelwerte in einer einzelnen Lehrveranstaltung nur eingeschränkt sinnvoll. Gegensätzliche Einschätzungen, daß eine Veranstaltung z.B. gut vorbereitet und anregend - oder nachlässig präpariert und langweilig sei, verlieren bei einer Einebnung ihren Wert für eine Auswertung. Auch andere Merkmale unterschiedlicher Typen des Lernens sollten in ihrem Profil zu erkennen sein. Die individuellen Einschätzungen und Präferenzen in Lerntempo, Zugang zum Stoff etc. müssen voll erhalten bleiben.

Für eine Diskussion der Ergebnisse mit den Studierenden kann es nützlich sein, die einzelnen Fragebögen nach der Auswertung so wieder auszugeben, daß die Einzelnen einen fremden Bogen vor sich haben. Die Demonstration gleicher oder gerade abweichender Sichtweisen ist für die Studierenden eine zusätzliche Lernchance.

6 Initiativen zur Evaluation der Lehre in Deutschland

Auslöser für praktische Schritte der Hochschulen, die über die Initiative einzelner Lehrender hinausgingen, waren die Ergebnisse der berühmt-berühmten Spiegel-Umfrage. Die TU Berlin als erste, einige Zeit später auch die Universitäten FU Berlin, Bielefeld, Bochum, Hannover und Stuttgart zogen die Konsequenz, sich ein Bild vom Zustand der Lehre durch Evaluation der Lehrveranstaltungen in möglichst vielen Studiengängen zu verschaffen. Ein weiterer Auslöser war auch die Veröffentlichung der durchschnittlichen lokalen, studiengangsspezifischen Studienzeiten durch den Wissenschaftsrat und andere Stellen. In Nordrhein-Westfalen löste ein entsprechender Erlaß des Ministeriums für Wissenschaft und Forschung zur Erprobung studentischer Veranstaltungskritik in einzelnen Studiengängen (nach anfänglichem Widerstand der Hochschulen) eine größere Anzahl von Erprobungen aus, z.B. an einer ganzen Reihe von Fachbereichen der Gesamthochschule Paderborn oder dem Fachbereich Wirtschaft der Fachhochschule Aachen sowie in 10 von 13 Fakultäten der Universität Bielefeld. Die HIS GmbH Hannover legte im Februar und Juli 1992 eine umfangreiche Dokumentation über (oft studentische) Initiativen zur Rückmeldung des Lehr- und Lernerfolgs vor, die auf großes Interesse in den Hochschulen stieß. Die Universität Göttingen und die TU Braunschweig haben gerade die Einführung der Lehrevaluation in allen Studiengängen beschlossen. Die Universitäten Mannheim, Konstanz, Gießen und Kaiserslautern beraten ebenfalls über die Lehrevaluation bzw. stehen vor dem Beginn. Darüber hinaus findet studentische Veranstaltungskritik an einer großen Zahl von Fachbereichen statt, ohne daß es allgemeine Beschlüsse an den betreffenden Hochschulen gibt.

Das Land Baden-Württemberg hat Mittel für Modellversuche zur Erprobung der "Befragung von Studierenden zur Bewertung der Lehre" bereitgestellt. Das BMBW hat zwei große überregionale Lehrevaluationen (Initiativen des RCDS und der LHG) gefördert, durch die ebenfalls Erfahrungen mit der Methodik, Organisation, Auswertung und Akzeptanz gesammelt werden sollen.

Die immer breitere Einführung der Lehrevaluation führt dazu, Defizite der Lehre zu lokalisieren und dabei nach längerer Zeit die Bedeutung der Studierenden als Zielgruppe der Lehre und Gegenstand der Verantwortung der Fachbereiche (neben der Selbstverantwortung der Studierenden) wieder einmal nachdrücklich ins Bewußtsein zu rücken.

7 Bisherige Erfahrungen

Im Wintersemester 1991/92 und Sommersemester 1992 sind in Nordrhein-Westfalen an vielen Stellen Veranstaltungsfragebögen zur Evaluation der Lehre entwickelt worden. Wenn auch durch massiven Anstoß von außen, durch das Ministerium für Wissenschaft und Forschung NRW, so ist doch an eine hochschuldidaktische Aktivität wieder angeknüpft worden, die in den 70er Jahren zunächst an Verbreitung gewann, dann aber eher im Verborgenen auf der Ebene individueller Lehrevaluation durch einzelne Lehrende weiterblühte. Die Aktion des Ministeriums für Wissenschaft und Forschung NRW zur Förderung studentischer Veranstaltungskritik hat eine ähnliche Wirkung gehabt wie die viel gescholtene Spiegel-Umfrage: Sie war - vor allem in dem vorgeschlagenen Fragebogen und der vermeintlichen Trennung von didaktischen und fachlichen Fragen - methodisch höchst angreifbar, hatte aber eine erhebliche politische Wirkung. Überdies mobilisierte sie viele Kräfte, die - wenn schon - dann eine solche Befragung auch auf "vernünftige" methodische Füße stellen wollten. Das ist erfreulich. Allerdings muß bei Durchsicht der Fragebögen doch bei vielen angemerkt werden, daß man nicht einerseits den Fragebogen-Vorschlag des Ministeriums für Wissenschaft und Forschung NRW methodisch kritisieren und andererseits dann fröhlich einen Fragebogen selbst stricken kann, ohne in der empirischen Sozialforschung zu Hause zu sein.

8 Fallbeispiel Bielefeld

Am Beispiel der Universität Bielefeld möchte ich einige Strategien und Instrumente der Lehrevaluation und Erstellung von Lehrberichten erläutern. Grundlage sind die Erfahrungen mit Lehrevaluation, die das IZHD Bielefeld seit 18 Jahren gesammelt hat, die flächendeckende praktische Durchführung und Auswertung in 10 von 13 Fakultäten, die Erstellung von Lehrberichten in ebenfalls 10 von 13 Fakultäten und ein BMBW-Projekt "Selbstevaluation von Fachbereichen (als Grundlage von Lehrberichten)", das ich zur Zeit durchführe.

Häufig knüpft die studentische Lehrbeurteilung an vorausgegangene Aktivitäten der Fachschaften an. Während sich aber solche Initiativen in der Regel daran orientieren, die Ergebnisse in "Hit-Listen" fakultätsöffentlich zu machen (und zum Teil die besten HochschullehrerInnen mit einem Preis auszuzeichnen), orientiert sich die Aktion in Bielefeld an einem hochschuldidaktischen Diskursmodell. Ziel der studentischen Lehrbeurteilung ist es dort, die Datengrundlage für einen Dialog zwischen Lehrenden und Lernenden zu gewinnen und damit die Lehre zu verbessern.

Die Auswertung traditioneller, handvercodeter Bögen führt vielfach dazu, daß die Ergebnisse viel zu spät vorliegen, um sie noch im gleichen Semester besprechen zu können. Spätere Auswertungen finden die betreffende Veranstaltung und ihre TeilnehmerInnen nicht mehr vor und sind auch für die Lehrenden nur noch von eingeschränktem Informationswert. Aus hochschuldidaktischer Sicht ist es unabdingbar, daß die Ergebnisse noch im gleichen Semester für ein Gespräch zwischen Lehrenden und Lernenden der beurteilten Veranstaltung genutzt werden. Geht eine Befragung über eine einzelne Lehrveranstaltung hinaus, die natürlich schnell ausgewertet werden könnte, muß der Fragebogen maschinenlesbar gestaltet werden, damit die Ergebnisse auch bei großen Teilnehmerzahlen noch im gleichen Semester zur Verfügung stehen. Das entsprechende Lesegerät, das auch gemietet werden kann, liest bis zu 3.000 Belege pro Stunde. Die Ergebnisse können daher bereits in der folgenden Lehrveranstaltung besprochen werden.

Die an der Universität Bielefeld entwickelten maschinenlesbaren Fragebögen mit ihren Kostenvorteilen und ihren Einsparungen an Auswertungszeit (binnen 8 Tagen) haben inzwischen Nachfragen aus zahlreichen anderen Hochschulen nach sich gezogen. Aus datenschutzrechtlichen und gesetzlichen Gründen ist die Beteiligung der einzelnen Fakultäten und ihrer Dozenten völlig freiwillig. Trotzdem hat sich in Bielefeld in allen 10 Fakultäten niemand ausgeschlossen. Überdies ist - als eine für die Beteiligung vielfach entscheidende vertrauensbildende Maßnahme - Vertraulichkeit gegenüber einer weiteren Hochschulöffentlichkeit zugesichert. Für die Auswertung der Daten sind die Fakultäten verantwortlich, und die Daten bleiben beim Dekan der jeweiligen Fakultät. Die VeranstalterInnen erhalten jeweils die Grundauszählung für ihre eigene Veranstaltung - mit der Erwartung, daß die Ergebnisse in den Veranstaltungen gemeinsam mit den Studierenden interpretiert und diskutiert werden. In einzelnen Fakultäten hatten sich die Lehrenden auch auf den Vergleich ihrer Ergebnisse geeinigt. Teilweise können auch die Fachschaften mit Genehmigung des Dekans die Ergebnisse auswerten. In einzelnen Fakultäten haben die Ausschüsse für Lehre die Einzelergebnisse durchgesehen und über Schlußfolgerungen für den Lehrbetrieb beraten. Dies war jedoch nur im Einvernehmen mit den Veranstaltern möglich.

Die Ergebnisse der bisherigen Lehrbeurteilungen haben bei vielen Lehrenden verbreitete Bedenken gegen die studentische Veranstaltungskritik abgebaut. Die eingesetzten, ausführlichen Fragebögen wurden im allgemeinen als sachgerecht und fair beurteilt. Selbst Skeptiker, die eine derartige Evaluation für überflüssig gehalten hatten, mußten einräumen, daß sie aus den Ergebnissen wesentlich mehr Nutzen ziehen konnten, als ursprünglich erwartet. Niemand ist angeprangert worden, wie vielfach vorher befürchtet. Die

Ergebnisse fielen im Gegenteil erheblich positiver aus, als die Lehrenden ursprünglich erwartet hatten. Das bedeutet nicht, daß nicht auch erhebliche Differenzen in der Lehrkompetenz hervortraten und auch insgesamt positiv Beurteilte auf Schwachpunkte oder sonstige wesentliche Umstände hingewiesen worden wären. Auf Grund dieser Erfahrungen besteht eine verbreitete Bereitschaft zu einer Befragung in jedem Semester. Davon ist allerdings bei einer Evaluation in allen Veranstaltungen wegen des zu erwartenden Abstumpfungseffektes bei den Studierenden abzuraten. Kürzere Befragungen könnten in jedem Studienjahr einmal, umfangreichere Erhebungen in einem drei- oder vierjährigen Rhythmus stattfinden.

9 Empirische Grundlagen der Lehrberichte

Über die Einführung studentischer Veranstaltungskritik hinaus haben 10 der 13 Fakultäten der Universität Bielefeld beschlossen, einen Situationsbericht zur Lage von Lehre und Studium ("Lehrbericht") zu erproben. Als Grundlage der Lehrberichte dienen

- eine Befragung der Lehrenden
 - zur Situation von Lehre und Studium,
 - zum Selbstkonzept als HochschullehrerIn,
 - zu ihrer Meinung über Studierende und Studienverhalten,
- eine Sonderauswertung der studentischen Veranstaltungskritik auf aggregierter Datenebene,
- eine Befragung der Studierenden zur Situation im Studium,
- Experteninterviews bzw. eine Gruppendiskussion mit dem Ausschuß für Lehre,
- Experteninterviews mit dem Prüfungsamt,
- zahlreiche Statistiken des Dezernats für Planung und Statistik,
- die Zielvorstellungen des Fachbereichs zu Lehre und Studium,
- eine Stellungnahme der Fachschaft zur Situation von Lehre und Studium (um den Vollständigkeits- und Realitätsgehalt des Gesamtberichts zu fördern).

Darüber hinaus wird daran gearbeitet, neben quantitativen Parametern auch die Qualität der vermittelten/erworbenen Kompetenz im Studium zu erfassen, u.a. deswegen, weil verkürzten Effizienzvergleichen ein Riegel vorge-schoben und reinen Kostengesichtpunkten die "Güte der Ausbildung" gegen-übergestellt werden soll.

Aus diesem Material wird ein selbstreflexiver, selbstkritischer Text erstellt, der in dieser Form wohl im Fachbereich verbleibt. Eine gemilderte Fassung

wird dann veröffentlicht, wobei ich darauf setze, daß die beiden Versionen sich durch Reformen des Fachbereichs immer mehr annähern.

10 Einige Schlußfolgerungen

1. Lehrberichterstattung und insbesondere studentische Veranstaltungskritik sind wichtige Mittel, Defizite der Lehre zu lokalisieren. Sie rücken nach längerer Zeit die Bedeutung der Studierenden als Zielgruppe der Lehre und Gegenstand der Verantwortung der Fachbereiche (neben der Selbstverantwortung der Studierenden) wieder einmal nachdrücklich ins Bewußtsein.
2. Lehrberichterstattung und Veranstaltungskritik stellen neue Chancen der Mitbestimmung und Kontrolle der Betroffenen dar, die wirksamer sein kann als Gremienarbeit,
 - a) weil sie unmittelbar auf die Inhalte und Formen des Studiums gerichtet ist und diese hinterfragen kann,
 - b) weil sie sich direkt an die verantwortlichen Lehrenden wendet, ohne ferne Mehrheitsbeschlüsse und Kompromisse in Gremien, die oft nicht einmal durchgesetzt werden,
 - c) weil sie nicht über einen Stimmenproporz verfälscht ist, sondern die Lehrenden direkt mit den Studierenden konfrontiert, so daß sie ihr Lehrangebot unmittelbar legitimieren müssen. De facto kommt hier keine Konfrontation, sondern kommen Aushandlungsprozesse in Gang, die geeignet sind, den Studierenden einen Teil ihrer Rechte zur Mitwirkung an der Studiengestaltung (ihrer Studiengestaltung) zurückzugeben.
3. Die Lehrberichterstattung kann, wenn sie ihren selbstreflexiven und selbstkritischen Ansatz beibehält und nicht zum Hochschulmarketing denaturiert, das geeignete Instrument zum Einstieg in die inhaltliche Studienreform sein.

Für Evaluationsvorhaben nützliche Literatur:

Anderson, S.; Ball, S.; Murphy, R.: Encyclopedia of Educational Evaluation. San Francisco 1976.

Aubrecht, J.D.: Reliability, Validity and Generalizability of Student Ratings of Instruction. ERIC Document Reproduction Service No. ED 213 296. Manhattan 1981. (IDEA Paper No. 6).

Bloom, B.; Hastings, G.; Madaus, G.F. (eds.): Handbook on Formative and Summative Evaluation of Student Learning. New York 1971.

Bülow, M.: Evaluierung von Projektstudiengängen als Aktionsforschung. Hamburg 1976. (Hochschuldidaktische Forschungsberichte 3).

Bülow, M.: Evaluation I. Verfahren, Methoden, Erfahrungen zur Überprüfung universitärer Ausbildung. Hamburg 1977. (Hochschuldidaktische Arbeitspapiere 8).

Cashin, W.E.: Student Ratings of Teaching. A Summary of the Research. Manhattan 1988. (IDEA Paper No. 20).

Cashin, W.E.; Noma, A.; Hanna, G.S.: Comparative Data by Academic Field. Manhattan 1987. (IDEA technical report No. 6).

Craft, A. (ed.): Quality Assurance in Higher Education. Proceedings of an International Conference. London/Washington DC 1992. (The Falmer Press).

Doyle, K.O.: Student Evaluation of Instruction. Lexington/Mass. 1975.

Fittkau, B. u.a.: Kommunizieren lernen (und umlernen). Braunschweig 1977.

Flood-Page, C.: Student Evaluation of Teaching. The American Experience. London 1974.

Harnier, L.v.; Schneider-Amos, I.: Auswirkungen einer Berufsausbildung auf das Studium der Betriebswirtschaftslehre. München 1992. (Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung 30).

Herbst, J.: Lehrleistungen als Kriterium des Hochschulwettbewerbs - die US-amerikanische Erfahrung. In: Webler, W.-D.; Otto, H.-U. (Hrsg.): Der Ort der Lehre in der Hochschule. Lehrleistungen, Prestige und Hochschulwettbewerb. Weinheim 1991.

Holtkamp, R.; Schnitzer, K. (Hrsg.): Evaluation des Lehrens und Lernens. Ansätze, Methoden, Instrumente. Evaluationspraxis in den USA, Großbritannien und den Niederlanden. Hannover 1992. (Hochschulplanung 92).

Huber, L.; Bürmann, I.; Francke, R.; Schmidt, W. (Hrsg.): Auswertung. Rückmeldung. Kritik im Hochschulunterricht. Hamburg 1978. (Blickpunkt Hochschuldidaktik 51).

Keil, W.; Piontkowski, U.: Strukturen und Prozesse im Hochschulunterricht. Weinheim/Basel 1973.

Luber, B.: Evaluation von verbalen Gruppenprozessen durch Beobachten - ein praktisches Beispiel. In: AUE-Informationen 1977, Sonderheft 21, S. 38 -53.

- Marsh, H.W.: Students' Evaluations of University Teaching. Dimensionality, Reliability, Validity, Potential Biases, and Utility. In: *Journal of Educational Psychology* 1984, 76, pp. 707 - 754.
- Ders.: Students' Evaluations of University Teaching. Research Findings, Methodological Issues, and Directions for Future Research. In: *International Journal of Educational Research* 1987, 11, pp. 253 - 388.
- McKeachie, W.J.: The New Reporters: Students Rate Instruction. In: Pace, C.R. (Ed.): *Evaluating Learning and Teaching*. San Francisco 1979. (New Directions for Higher Education 4), pp. 59 - 75.
- Meyer-Althoff, M.: *Evaluation II. Versuche und Erfahrungen mit Selbstevaluation*. Hamburg 1978. (Hochschuldidaktische Arbeitspapiere 11).
- Miller, R.I.: *Evaluating Faculty for Promotion and Tenure*. San Francisco 1987.
- Müller-Wolf, H.M.; Fittkau, B.: Lehrverhalten von Hochschullehrern und seine Bedeutung für Einstellung und Verhalten von Studenten. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 1971, 3, S. 165 - 180.
- Parlett, M; Dearden, G. (eds.): *Introduction to Illuminative Evaluation*. Studies in Higher Education. 2nd. ed. Guildford (U.K.) 1981.
- Richter, R.: Qualitätsevaluation von Lehre und Forschung an den Universitäten der Niederlande. In: Webler, W.-D.; Otto, H.-U. (Hrsg.): *Der Ort der Lehre in der Hochschule. Lehrleistungen, Prestige und Hochschulwettbewerb*. Weinheim 1991. (Blickpunkt Hochschuldidaktik 90).
- Rossi, P.H.; Freemann, H.E.; Wright, S.R.: *Evaluation, a Systematic Approach*. London 1979.
- Schott, E.: *Zur empirischen und theoretischen Grundlegung eines Bewertungsinstrumentes für Vorlesungen*. Hamburg 1973. (Blickpunkt Hochschuldidaktik 28).
- Seldin, P.: *How Colleges Evaluate Professors. Current Policies and Practices in Evaluating Classroom Teaching Performance in Liberal Arts Colleges*. New York 1975.
- Ders.: *Successful Faculty Evaluation Programs. A Practical Guide to Improve Faculty Performance and Promotion/Tenure Decisions*. Crugers, NY 1980.
- Struening, E.L; Guttentag, M.: *Handbook of Evaluation Research. I*. Beverly Hills/London 1975.
- Tiefgens, H. u.a.: *Aufgaben und Probleme der Evaluation in der Erwachsenenbildung*. Bonn 1986.
- Webler, W.-D.: Forschungsansätze in der offenen Entwicklung und Evaluation sozialer Experimentalprogramme. In: Hellstern, D.-M.; Wollmann, H. (Hrsg.): *Experimentelle Politik - Reformstrohfeuer oder Lernstrategie. Bestandsaufnahme und Evaluierung*. Opladen 1983.
- Ders.: *Evaluation soziologischer Lehrveranstaltungen. - Gutachten für den Vorstand der Deutschen Gesellschaft für Soziologie*. (Unv. Man.) Bielefeld 1989.

Webler, W.-D.; Otto, H.-U. (Hrsg.): Der Ort der Lehre in der Hochschule. Lehrleistungen, Prestige und Hochschulwettbewerb. Weinheim 1991. (Blickpunkt Hochschuldidaktik 90).

Winteler, A.: Schätzverfahren. Ein Bericht über anglo-amerikanische Untersuchungen. In: Huber, L. u.a. (Hrsg.): Auswertung. Rückmeldung. Kritik im Hochschulunterricht. Hamburg 1978.

Zentrum für päd. Berufspraxis (Hrsg.): Evaluation zur Auswertung von Projekten, Lehrveranstaltungen und berufspraktischen Vorhaben. Oldenburg 1977.

Anschrift des Verfassers:

Dr. Wolff-Dietrich Webler
Interdisziplinäres Zentrum für
Hochschuldidaktik
Universität Bielefeld
Universitätsstraße 25

33615 Bielefeld