

Rainer Schnell und Johannes Kopp

Zur Evaluation von Lehrveranstaltungen in den Sozialwissenschaften

In deutlicher Verkürzung dessen, was unter „Evaluation der Lehre“ verstanden werden kann, erfolgt in der Diskussion um die Bewertung der Leistungen der Hochschulen zumeist eine Reduktion auf die Abgabe von Urteilen der Studierenden in Hinsicht auf ihre Einschätzung der jeweiligen Lehrveranstaltungen. Erstaunlicherweise wird die Diskussion um die Möglichkeiten und Probleme studentischer Evaluationen unter häufiger Missachtung sowohl methodischer Standards der Evaluationsforschung als auch des Ausmaßes der tatsächlichen Durchführung, der verwendeten Instrumente und der Kooperationsbereitschaft der beteiligten Akteure geführt.

Um die Diskussion über die Evaluation der Lehre in den Sozialwissenschaften auf eine empirische Basis zu stellen, haben wir im Rahmen eines empirischen Forschungsprojekts an der Fakultät für Verwaltungswissenschaft der Universität Konstanz zunächst versucht, einen Überblick über die Verbreitung studentischer Evaluationen durch eine Befragung der sozialwissenschaftlichen Fakultäten in der Bundesrepublik zu gewinnen. Dieser Beitrag stellt die Ergebnisse dieser Befragung vor. Abschließend präsentieren wir neue methodologische Vorschläge zur Durchführung von Lehrevaluationen.¹

1 Befragung bundesdeutscher Hochschulen zu studentischen Lehrevaluationen in sozialwissenschaftlichen Studiengängen

Um die Frage nach der Verbreitung derartiger studentischer Lehrbefragungen im Rahmen sozialwissenschaftlicher Studiengänge an den Hochschulen der Bundesrepublik Deutschland zu beantworten, wurde eine Befragung derjenigen Institute, Fachbereiche oder Fakultäten durchgeführt, an denen ein Studium der Sozialwissenschaften beziehungsweise eines ihrer Teilfächer möglich ist. Entsprechend wurden die Adressen aller universitären Einrichtungen ermittelt, die die Fächer Sozialwissenschaften, Soziologie, Politikwissenschaften sowie Ver-

1 Der vorliegende Bericht entstand im Zusammenhang eines durch das Ministerium für Wissenschaft, Forschung und Kunst des Landes Baden-Württemberg geförderten Forschungsprojektes „Fakultätsinterne Evaluation der Lehre – die Weiterentwicklung des bisherigen Evaluationskonzepts“ (vgl. Schnell/Kopp 2000).

waltungswissenschaft als Studienfach anbieten. Als Grundlage der entsprechenden Adressensammlung diente der bundesweite Internet-Studienführer, der von der zentralen Studienberatung der Universität Münster organisiert wird. Insgesamt 94 verschiedene Adressaten ergaben sich als Ergebnis dieser Recherche.²

Nicht immer konnte dabei eindeutig eine für die Durchführung der sozialwissenschaftlichen Studiengänge federführende Institution ausgemacht werden. Teilweise wurden aus diesem Grunde mehrere Fachgebiete umfassende Fachbereiche oder Fakultäten angeschrieben. Die entsprechenden Antworten beziehen sich aus diesem Grunde in dem ein oder anderen Fall also auf umfassendere Einheiten als die einzelnen Studiengänge der Sozialwissenschaften. In den meisten Fällen war es jedoch möglich, die für die Fächer Soziologie und Politikwissenschaften verantwortlichen Einheiten direkt zu ermitteln und anzuschreiben. Trotzdem sollte beachtet werden, dass die Befragungseinheiten sehr heterogene Organisationen mit unterschiedlichsten Größen der vertretenen Fächer, sowohl in Hinsicht auf die Zahl der Studierenden als auch in Hinsicht auf die Zahl der Lehrenden, darstellen.

2 Durchführung der Befragung

Für die Befragung wurde ein nur eine Seite umfassender standardisierter Fragebogen entwickelt, der zusammen mit einem Anschreiben an die Befragungseinheiten versandt wurde. Die Adressaten wurden gebeten, neben dem Fragebogen die in ihrer Institution eingesetzten Befragungsinstrumente zur Lehrevaluation sowie eventuell vorhandene weitere Dokumente zur internen Lehrevaluation zurückzusenden. Die Versendung der entsprechenden Anfragen erfolgte Anfang November 1999, ein Erinnerungsschreiben wurde nach drei Wochen im Dezember 1999 verschickt. Von den angeschriebenen 94 Fachbereichen, Fakultäten und Instituten antworteten insgesamt 81 auf unsere Anfrage. Die Antwortquote von 86,2 Prozent liegt deutlich über dem sonst bei schriftlichen Befragungen zu erwartenden Maße und kann wohl als ein Indiz für das Interesse der Adressaten für derartige Themen interpretiert werden.

3 Ausmaß der Verbreitung studentischer Lehrevaluationen

Von den an dieser Befragung teilnehmenden 81 Fachbereichen, Instituten oder Fakultäten mit sozialwissenschaftlichen Studiengängen haben mehr als zwei Drittel bereits mindestens einmal eine Lehrevaluation durchgeführt (vgl. Tabelle 1).³

-
- 2 Unter „sozialwissenschaftlichen“ Studienanschlüssen werden hier Studienabschlüsse in Soziologie, Politikwissenschaften und Verwaltungswissenschaften verstanden. Eine Berücksichtigung der Evaluationen in anderen, sicherlich ebenfalls den Sozialwissenschaften zurechenbaren Teildisziplinen (z. B. Sozialpsychologie, Ökonomie) erfolgte nicht, da dies den finanziellen Rahmen des Projekts gesprengt hätte.
 - 3 Zwei Institute reagierten zwar auf die Anfrage, beantworteten jedoch nicht den beigegeführten Fragebogen. In einem der beiden Fälle geschah dies mit dem Hinweis auf die sehr un-

Tabelle 1: Erfahrungen mit Lehrevaluation

| | Anzahl | Anteil in Prozent |
|--|--------|-------------------|
| Lehrevaluation in jedem Semester | 26 | 32,9 |
| Lehrevaluation schon durchgeführt, aber nicht jedes Semester | 30 | 38,0 |
| Noch keine Lehrevaluation durchgeführt | 23 | 29,1 |

Interessant ist nun jedoch auch, in welchem Umfang die entsprechenden Evaluationen durchgeführt werden. Hierzu wurde sowohl untersucht, ob alle oder nur bestimmte Personengruppen innerhalb des Faches, als auch, ob alle oder nur bestimmte Veranstaltungstypen evaluiert werden.

Tabelle 2: Lehrevaluation bei einzelnen Gruppen (n = 54)

| | Ja (in Prozent) |
|---|-----------------|
| Lehrevaluation bei Professoren | 96,3 |
| Lehrevaluation bei Mitarbeitern (Mittelbau) | 96,3 |
| Lehrevaluation bei Tutoren | 77,7 |

In insgesamt drei Viertel aller Evaluation betreibenden Institutionen werden alle bei der Lehre Beteiligten – also Professoren, der Mittelbau wie auch Tutoren – bewertet (vgl. Tabelle 2). Wenn einzelne Teilgruppen von der Evaluation ausgenommen werden, so sind dies vor allem die Tutoren. Diese werden jedoch immer noch in mehr als 77 Prozent aller Universitäten evaluiert. Auch hinsichtlich der Bewertung einzelner Lehrveranstaltungstypen findet sich ein ähnliches Bild. In fast 70 Prozent aller hier befragten Institutionen werden bei Evaluationen alle Arten von Lehrveranstaltungen bewertet, allerdings finden sich auch hier typische Unterschiede (vgl. Tabelle 3).

Tabelle 3: Lehrevaluation bei einzelnen Veranstaltungstypen (n = 54)

| | Ja (in Prozent) |
|--------------------------------------|-----------------|
| Lehrevaluation bei Vorlesungen | 94,4 |
| Lehrevaluation bei Seminaren/Übungen | 92,6 |
| Lehrevaluation bei Tutorien | 75,9 |

Wie nach den zuvor zitierten Ergebnissen hinsichtlich der einzelnen universitären Gruppen erwartbar, so werden vor allem die Vorlesungen und – zu einem nur unbedeutend geringeren Anteil – Seminare und Übungen bewertet.

terschiedliche Handhabung der Evaluation innerhalb der Fakultät. Freundlicherweise wurde jedoch das dabei teilweise zum Einsatz kommende Befragungsinstrument beigelegt, so dass es hier Berücksichtigung finden kann. Für die in diesem Abschnitt des Textes vorgenommenen Analysen vermindert sich jedoch die Stichprobe auf 79 Einheiten. Bei zwei Fragebögen wurde zwar angegeben, dass Evaluationen stattfinden, die weiteren Fragen jedoch nicht beantwortet. Aus diesem Grunde verringert sich die Stichprobe für die Tabellen 2 und 3 auf 54.

4 Überblick über die eingesetzten Befragungsinstrumente

Von den antwortenden 81 Instituten, Fachbereichen oder Fakultäten wurden insgesamt 38 Fragebögen zurückgesandt. Zwei weitere Institutionen verwiesen auf die Dokumentation ihrer Evaluationsbemühungen im Internet, aber immerhin 17 der Institutionen, die angaben, bereits einmal eine Evaluation durchgeführt zu haben, haben trotz der deutlichen Bitte keinen entsprechenden Bogen beigelegt.⁴ Hierbei besteht kein statistisch bedeutsamer Zusammenhang zwischen der Regelmäßigkeit der Lehrevaluation und der Bereitschaft, einen Fragebogen zurückzusenden.

Schon ein kurzer Blick auf die verschiedenen Evaluationsinstrumente zeigt die große Unterschiedlichkeit im Umfang der einzelnen Instrumente. Während einige Befragungsinstrumente bei einer sorgfältigen Bearbeitung durch die Studierenden sicherlich einen hohen Zeitaufwand mit sich bringen, umfassen andere Evaluationsbögen nur wenige Statements. Weiterhin handelt es sich nicht bei allen Materialien um fertige Befragungsinstrumente; zum Teil sind es Itemsammlungen, bei denen aus einer Vielzahl von Bewertungsdimensionen ein eigenes Instrument zusammengestellt werden kann. Zusammenfassend lassen sich bei den Evaluationsbögen zwei Idealtypen unterscheiden:

- Auf der einen Seite steht dabei die kurze und im Rahmen üblicher Veranstaltungen wohl recht problemlos durchzuführende Befragung, deren Instrument meist auf einer einzigen Seite zusammengefasst ist.
- Auf der anderen Seite befinden sich Instrumente, die eigentlich nicht der Evaluation einzelner Veranstaltungen, sondern einer erweiterten Analyse des Studierendenverhaltens und der allgemeinen Studienbedingungen sowie dem Studienablauf dienen sollen und die aus diesem Grunde auch wesentlich komplexer und vor allem auch umfangreicher gestaltet sind.

Trotz dieser Heterogenität des zugrundeliegenden Materials lassen sich eine Reihe von gemeinsamen Fragedimensionen benennen, die sich in den verwendeten Evaluationsinstrumenten finden. Dabei erscheint eine Unterscheidung von 11 Dimensionen sinnvoll (vgl. Tabelle 4). Es muss betont werden, dass sich hinter einer Nennung unterschiedlich umfangreiche und auch unterschiedlich sinnvolle Fragen und Operationalisierungen verbergen.

Es zeigen sich hier interessante Unterschiede zwischen den Instrumenten. Fast immer findet sich eine Einschätzung zur Konzeption und Struktur der Lehrveranstaltung oder zur Vorgehensweise der Dozierenden. Seltener – und vor allem eher bei Untersuchungen zur allgemeinen Studienmotivation als bei konkreten Veranstaltungsbewertungen – finden sich ausführlichere Fragen zum sozialen Hintergrund und anderen demographischen Variablen.

⁴ Insgesamt 23 Institute hatten noch keine Evaluationserfahrung, ein weiteres Institut reagierte zwar auf unser Anschreiben, füllte jedoch weder den Fragebogen aus noch wurde ein Evaluationsbogen beigelegt.

Tabelle 4: Berücksichtigung der einzelnen Dimensionen in den Instrumenten (n=38)

| Dimension | Anteil in Prozent |
|--|--------------------------|
| Konzeption und Struktur der Lehrveranstaltung | 84,2 |
| Vorgehensweise der Dozierenden | 84,2 |
| Welche Veranstaltung | 71,1 |
| Demographische Kurzangaben | 65,8 |
| Möglichkeit zur offenen Bewertung | 63,2 |
| Engagement der Studierenden | 60,5 |
| Art der Interaktionsbeziehungen in den Sitzungen | 31,6 |
| Gesamturteil | 28,9 |
| Ausführlichere demographische Angaben | 28,9 |
| Rahmenbedingungen | 23,7 |
| Motivation zum Besuch der Lehrveranstaltung | 21,1 |

5 Methodische Standards der Instrumente

Ein gründlicherer Blick auf die Instrumente macht zudem deutlich, dass viele der hier zu findenden Fragebögen recht konkret auf die Situation in den einzelnen Fachbereichen, Instituten oder Fakultäten bezogen sind. Entsprechend finden sich in vielen Instrumenten die üblichen Probleme eines nicht sorgfältig in vielen Pretests geprüften Instruments: Logische Widersprüche, Unklarheiten, Probleme mit der Filterführung, undefinierte und fälschlich als bekannt vorausgesetzte Begriffe. Man kann sich kaum des Eindrucks entziehen, dass die allermeisten dieser Befragungsinstrumente nicht entsprechend den methodischen Standards der empirischen Sozialforschung getestet worden sind. Wenn man Lehrevaluationen aber ernstnehmen will, dann müssen die Mindestanforderungen an sozialwissenschaftliche Instrumente natürlich auch hier erfüllt sein (vgl. beispielsweise dagegen für die Niederlande: Willems/Gijsselaers/de Bie 1994: 60). Interessanterweise existiert eine anscheinend in der Praxis der Lehrevaluation kaum rezipierte Itemsammlung, das sogenannte „Heidelberger Inventar zur Lehrveranstaltungs-Evaluation“, das zahlreiche Test- und Erprobungsphasen durchlief (vgl. hierzu ausführlich Rindermann 1996).⁵ Natürlich finden sich in diesem Inventar viele der oben erwähnten Dimensionen und ähnliche Operationalisierungen wie in den untersuchten Instrumenten. Aus diesem Grunde lässt sich durchaus vermuten, dass in der Regel auch mit den üblichen Instrumenten nicht reine Artefakte produziert werden. Trotzdem sind umfangreiche Tests und Validierungsmaßnahmen für jedes verwendete Instrument erforderlich, wenn ein Messanspruch erhoben werden soll – spätestens dann, wenn ein Teil der Ressourcenallokation über die Ergebnisse solcher Evaluationen erfolgen soll, wird dies (schon juristisch) unverzichtbar. Dies setzt aber einen nicht unerheblichen finanziellen und personellen Aufwand bei der Instrumentenkonstruktion voraus.

6 Teilnahmemotivation der Lehrenden und Lernenden

In der Diskussion um die Evaluation der Lehre wurde von Anfang an von Problemen mit der Teilnahmebereitschaft an entsprechenden Befragungen berichtet. Z. B. gibt Grünh (1992) für eine der ersten Lehrevaluationen dieser Art in der Bundesrepublik an, dass circa 80 Prozent aller Veranstaltungen und darin wiederum rund 70 Prozent aller Studierenden erfasst werden konnten. Bestenfalls wurden hier also nur etwa 56 Prozent aller Studierenden überhaupt erreicht. Die Ergebnisse der seit 1996 durchgeführten Lehrevaluation an der Universität Konstanz bestätigen diese Größenordnung.

Wenn man über die Gründe für diese doch nicht allzu große Teilnahmebereitschaft spekulieren will, so kann man auf studentischer Seite die meist völlige Konsequenzenlosigkeit einer entsprechenden Evaluation vermuten. In Hinsicht auf die Dozierenden kann ein Selektivitätseffekt angenommen werden: Wer eine schlechtere Bewertung befürchtet, kann diese durch Nichtteilnahme einfach vermeiden. Gerade vor dem Hintergrund einer unter Umständen höchst selektiven Teilnahmebereitschaft sollten die inhaltlichen Ergebnisse entsprechender Studien nur vorsichtig interpretiert werden.⁶

Um einen Eindruck von der Teilnahmebereitschaft an Lehrevaluationen an den Hochschulen trotz der vermuteten Abwesenheit systematischer Untersuchungen zu erhalten, wurde in der hier berichteten Befragung bundesdeutscher Hochschulen die Beurteilung der Teilnahmebereitschaft der Studierenden und der Lehrenden auf einer fünfstufigen Skala erfragt (vgl. Tabelle 5).⁷

Auf den ersten Blick scheint die Teilnahmebereitschaft sowohl der Studierenden wie auch der Lehrenden recht hoch zu sein. Der Median beider Variablen liegt bei 2, im arithmetischen Mittel wird die Teilnahmebereitschaft nur ein wenig schlechter als „gut“ eingeschätzt. Es gibt jedoch zu bedenken, dass in mehr als einem Viertel der Institute die Teilnahmebereitschaft der Studierenden höchstens als „befriedigend“ eingestuft wurde und diese Einschätzung in fast einem Drittel aller Fälle auch für die Teilnahmebereitschaft der Lehrenden gilt. Betrachtet man den Zusammenhang dieser beiden Einschätzungen, so finden sich in 60 Prozent der Institute keine Unterschiede zwischen Studenten und Lehrenden. Nur in 14 Prozent aller Fälle weicht die Einschätzung um mehr als einen Punkt voneinander ab.

6 Zusätzlich gilt es zu beachten, dass Lehrevaluationen meist am Ende einer Veranstaltung durchgeführt werden, so dass besonders unzufriedene Studierende – soweit es sich nicht um eine Pflichtveranstaltung handelt – wahrscheinlich nicht mehr teilnehmen.

7 Es ist an dieser Stelle daran zu erinnern, dass die jeweiligen Fragebögen wohl von sehr unterschiedlichen Personengruppen beantwortet wurden und deshalb sich auch auf sehr unterschiedliche Erfahrungswerte beziehen.

**Tabelle 5: Teilnahmebereitschaft an studentischen Lehrevaluationen
(in Prozent der jeweiligen Spalte)**

| | Teilnahmebereitschaft | |
|--------------|-----------------------|----------|
| | Studierende | Lehrende |
| sehr gut | 21,2 | 23,5 |
| gut | 51,9 | 43,1 |
| befriedigend | 15,4 | 29,4 |
| ausreichend | 7,7 | 3,9 |
| mangelhaft | 3,8 | 0,0 |
| n | 52 | 51 |

Wenn diese Ergebnisse die reale Situation bei der Evaluation von Lehrveranstaltungen durch Studierende wiedergeben, so müssen die entsprechenden Ergebnisse mit einer vielleicht noch größeren Vorsicht interpretiert werden. Da die verschiedenen Ausfallprozesse zudem theoretisch eher zu einer positiven Verzerrung führen – so fehlen besonders frustrierte Studierende häufiger bei der Evaluation, Lehrende mit einem eventuell begründeten schlechteren Selbstbild ihrer Lehre weisen wohl eine geringere Teilnahmebereitschaft auf – sind selbst positive Ergebnisse einer derartigen Evaluation kein sicheres Kriterium guter Lehre.

7 Vorschläge zur Durchführung von Lehrevaluationen

Eine grundlegende Erkenntnis der Literatur zur allgemeinen Evaluationsforschung ist, dass es ohne eine klare Festlegung der Ziele einer bestimmten Maßnahme letztlich nicht möglich ist, den Grad der Zielerreichung festzulegen. Gerade im Bereich der Universität aber erscheint eine Übereinkunft, was unter guter Lehre verstanden werden soll, nicht gegeben. Nicht zuletzt aus diesem Grunde können Evaluationen ab und an als ‚publikumswirksame Selbstprofilierung‘ (Brandstädter 1990: 224) oder gar als reiner Aktionismus erscheinen.

Will man im Bereich der Hochschulbildung tatsächlich Maßnahmen zur Qualitätssteigerung durchführen, dann sind methodisch anspruchsvolle Evaluationen unvermeidlich. Zwar spielen hierbei studentische Einschätzungen der Lehrqualität auch eine Rolle, diese müssen jedoch durch eine ganze Reihe anderer Maßnahmen ergänzt werden (vgl. hierzu ferner auch Schencker-Wicki 1996).

Wie schon dargelegt, müssen zunächst die verwendeten Evaluationsinstrumente den Standards der empirischen Sozialforschung entsprechen. Weiterhin muss jede Durchführung einer studentischen Lehrevaluation begleitet werden durch eine Untersuchung der nicht erfassten Veranstaltungen sowie der Veranstaltungsabbrecher und derjenigen Studenten, die zu keiner Lehrveranstaltung erscheinen. Ohne solche Maßnahmen können verschiedene Selektions- oder Auswahlprozesse zu einer vermutlich positiven Verzerrung der Ergebnisse führen.

Das zentrale Problem studentischer Lehrevaluationen besteht aber in der Validität der Messungen. Um diese beurteilen zu können, muss geklärt werden, was unter guter Lehre zu verstehen ist.⁸ Dies ist keine mit Hilfe methodischer Kriterien beantwortbare Frage. Vielleicht kann man sich pragmatisch auf „akademischen Erfolg der Teilnehmer“ einerseits und „kommerziellen Erfolg der Teilnehmer“ im Sinne des Arbeitsmarktes andererseits einigen. Beides ließe sich konzeptuell relativ leicht operationalisieren. Die Erhebung solcher Kriterien ist aber dann unverzichtbar, wenn Lehrevaluationen wirkliche Evaluationen und keine Meinungsumfragen darstellen sollen. Eine isolierte Verbesserung der Instrumente oder auch die Kontrolle der Selektivitätsprobleme wird kaum etwas zu einer Verbesserung der Evaluationen beitragen.

Die Verbesserung von Lehrevaluationen bedeutet also eine Ausweitung der durchzuführenden Maßnahmen. Konkret schlagen wir vor:

1. Es sollten neben der Befragung der Studierenden regelmäßige Befragungen der Absolventen – einschließlich aller Abbrecher – erfolgen. Der Katalog der Fragen sollte sich vor allem auf die Dauer bis zum Jobantritt und die üblichen Bewertungskriterien des Jobs beziehen.
2. Hierzu ist der Aufbau einer regelmäßigen – erfahrungsgemäß personalintensiven – Adressenpflege notwendig.
3. Zusätzlich sollten die akademischen Karrieren aller Doktoranden und Habilitanden – etwa über die Datenbanken SOLIS und SSCI – verfolgt werden. Deren Zitatshäufigkeiten sollten ebenfalls Bestandteil der Evaluation werden.
4. Um die methodischen Probleme unterschiedlichen Antwortverhaltens in verschiedenen Veranstaltungen kontrollieren zu können und vor allem um für die langfristige Evaluation der Veranstaltungen die Zuordnung der Studenten zu den Veranstaltungen verfügbar zu machen, ist die Verwendung einer eindeutigen Identifikationsnummer unverzichtbar. Aus rein pragmatischen Gründen empfiehlt sich die Verwendung der Matrikelnummer. Dies setzt voraus, dass die Studenten die Notwendigkeit dieser Maßnahme einsehen.
5. Die Beantwortung der Lehrevaluation sollte nicht mehr durch Paper-Pencil-Methoden erfolgen, sondern in Form eines Internet-Surveys.⁹ Nach Einloggen mit Matrikel und Passwort wird der Fragebogen bearbeitet. Dieses Vorgehen

8 Zu einer ähnlichen Einschätzung kommt Gold (1996: 149), wenn er schreibt: „Das Validieren eines Inventars zur Lehrqualität setzt eine theoretische Begründung der Kriterien guter Lehrveranstaltungen voraus. Der Validitätsdiskussion müsste also eine Kriteriendiskussion vorangehen“. Es bedarf jedoch eines gewissen Optimismus zu glauben, dass man hier zu einem einfachen und einheitlichen Ergebnis über die Kriterien guter Lehre gelangen kann.

9 Zwar gibt es einige Ansätze zur Online-Lehrevaluation (zum Beispiel an der Wirtschaftsuniversität Wien, der Universität Greifswald und der Technischen Universität Hamburg-Harburg), diese sind aber nicht Bestandteil eines umfassenden Lehrbewertungskonzepts, sondern lediglich die Online-Form einer herkömmlichen Befragung, die allenfalls um ein halbautomatisches Auswertungsmodul erweitert wurde.

bietet eine Reihe von Vorteilen: Druck- und Organisationskosten verschwinden, die Kooperation der Lehrenden ist nicht notwendig, die Ergebnisse liegen maschinenlesbar vor, die Auswertung kann automatisch erfolgen, eine studentische Responsequote ist berechenbar, die Evaluation kann gegebenenfalls mehrmals pro Semester durchgeführt werden und so schon während der Veranstaltung zu einer Korrektur führen. Die Entwicklungskosten für ein solches Systems dürften sich schon nach kurzer Zeit amortisieren. Der einzige Nachteil liegt in der unverzichtbaren Notwendigkeit eines Personenidentifiers, in der Regel der Matrikelnummer.

Man kann leicht sehen, dass die Durchführung dieser Maßnahmen weder kosten- noch personalneutral ist. Faktisch ist der Aufwand erheblich. Will man eine Bewertung der Lehre empirisch begründen, so ist dieser Aufwand aber unvermeidlich. Nimmt man den Aufwand in Kauf, dann gibt es nur noch ein faktisches Problem: die Kooperationsbereitschaft der Studenten. Kann man diese nicht von der Notwendigkeit der Verknüpfung ihrer Angaben mit den Daten der Absolventenbefragungen und der Daten der Prüfungsstatistik überzeugen, ist eine Validierung der Lehrevaluationen unmöglich. Sollten noch durchzuführende Voruntersuchungen zeigen, dass diese Voraussetzung nicht erfüllbar ist, so schlagen wir die Einstellung der bisherigen Praxis der rituellen Perpetuierung unvalidierter Meinungsbefragungen anstelle von Lehrevaluationen vor.

Literatur

- Brandstätter, Jochen, 1990: Evaluationsforschung: Probleme der wissenschaftlichen Bewertung von Interventions- und Reformprojekten. *Zeitschrift für Pädagogische Psychologie* 4: 215-227.
- Gold, Andreas, 1996: Können Studierende die Qualität der Lehre beurteilen? Einige Anmerkungen zu Rindermanns Antwort an seine Kritiker. *Zeitschrift für Pädagogische Psychologie* 10: 147-150.
- Grühn, Dieter, 1992: Freie Universität Berlin. Diskussion über Qualität der Lehre in Gang gekommen: Hohe Beteiligung an der Studentenbefragung. *Das Hochschulwesen* 40: 92-
- Rindermann, Heiner, 1996a: Untersuchungen zur Brauchbarkeit studentischer Lehrevaluationen. Landau: Verlag Empirische Pädagogik.
- Schencker-Wicki, Andrea, 1996: Evaluation von Hochschulleistungen. Leistungsindikatoren und Performance Measurements. Wiesbaden: Deutscher Universitätsverlag.
- Schnell, Rainer, Kopp, Johannes, 2000: Theoretische und methodische Diskussionen der Lehrevaluationsforschung und deren praktische Bedeutung. Endbericht des Forschungsprojektes „Fakultätsinterne Evaluation der Lehre – die Weiterentwicklung des bisherigen Evaluationskonzepts“. Konstanz (Volltext verfügbar unter „<http://www.ub.uni-konstanz.de/kops/volltexte/2001/605>“).
- Willems, J., Gijsselaers, W., de Bie, D., 1994: Qualitätssorge in der Lehre. Leitfaden für die studentische Lehrevaluation. Neuwied: Luchterhand.