

Methodik und Anwendung der Lehrveranstaltungsevaluation für die Qualitätsentwicklung an Hochschulen

Rindermann, Heiner

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Rindermann, H. (2003). Methodik und Anwendung der Lehrveranstaltungsevaluation für die Qualitätsentwicklung an Hochschulen. *Sozialwissenschaften und Berufspraxis*, 26(4), 401-413. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-38178>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. 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. 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:

This document is made available under Deposit Licence (No Redistribution - no modifications). 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. 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.

Methodik und Anwendung der Lehrveranstaltungsevaluation für die Qualitätsentwicklung an Hochschulen

Heiner Rindermann

Der Einsatz von Evaluation soll neben Erkenntnisgewinn (Beschreibung und Bewertung) positive Konsequenzen für die Praxis haben. An Hochschulen als Gegenstand und Ausführende von Evaluation sind Forschung und Praxis der Evaluation besonders eng verknüpft, ohne dass sich diese Konstellation aber bislang für beide Seiten besonders fruchtbar ausgewirkt hätte. Durch Lehrveranstaltungsevaluation sollte sich die Lehrqualität in Veranstaltungen verbessern, die Akzeptanz von Lehre erhöhen, Studierende sollten als Folge verbesserter Lehre über ein umfangreicheres, vertiefteres und anwendbareres Wissen verfügen. Ergebnisse der Lehrerevaluationsforschung zeigen indessen, dass Lehrqualität nicht automatisch durch Evaluation verbessert wird, hinzutreten müssen förderliche Kontextmerkmale und vor allem Maßnahmen, die Hilfestellung bei der Rückmeldung geben sowie Beratung und Training für Dozenten beinhalten. Möglichkeiten und Probleme bei der Verbindung wissenschaftlicher Methodik und Praxis an Hochschulen werden diskutiert.

1 Evaluation als angewandte sozialwissenschaftliche Methodik

Evaluation und *Evaluationsforschung* können auf eine lange Tradition zurückblicken, die wissenschaftlich unterstützte Bewertung von Eingriffen zählt in den verschiedensten Anwendungsfeldern zu den Kernaufgaben angewandter Forschung (Wotta/Thierau 1998; Beywl 1988). Als selbständiges Forschungsgebiet etablierte sich die Evaluationsforschung aber erst in den 1960er und 1970er Jahren in den USA infolge der damals aufgelegten großen Förderprogramme. Um der Bewilligung oder Fortsetzung von Maßnahmen eine objektive Basis zu geben, wurden Forscher beauftragt, diese Programme und ihre Effekte mit wissenschaftlichen Methoden zu untersuchen. Evaluation soll für politische, ökonomische oder administrative Entscheidungen mit

Methoden der Wissenschaft eine auf rationalem Wege ermittelte Informationsgrundlage bereitstellen und diese Entscheidungen so legitimieren. Lehrevaluation und Lehrevaluationsforschung sind aber schon älter, schon seit den 1920er Jahren werden diese an nordamerikanischen Universitäten praktiziert (Marsh 1987).

Unter *Evaluation* soll nun die systematische Analyse und empirische Untersuchung von Konzepten, Bedingungen, Prozessen und Wirkungen zielgerichteter Aktivitäten zum Zwecke ihrer Bewertung und Modifikation verstanden werden (Rossi et al. 1999). Gegenstände von Evaluationen sind Programme, Verfahren, Produkte, Institutionen und Tätigkeiten von Personen. Über adäquate *Beschreibung* und *Messung* hinausgehend *bewerten* Evaluationen und *optimieren*. Formative Ansätze legen hierbei den Schwerpunkt auf die Verbesserung von laufenden Verfahren, summative Ansätze nehmen über die Bereitstellung fundierter Informationen Wirkungsbeurteilungen vor und optimieren so Entscheidungen.

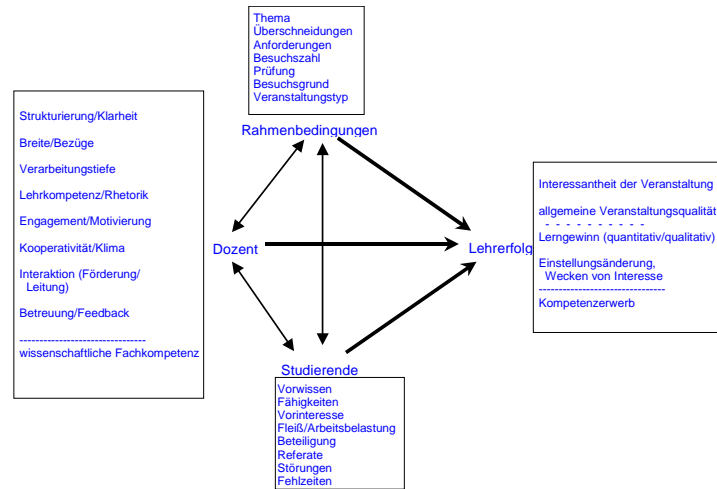
In Evaluation und Evaluationsforschung lassen sich verschiedene Modelltypen voneinander abheben, die Evaluationsmodelle unterscheiden sich hinsichtlich der Betonung versuchsplanerischer Methodik oder naturalistischer Designs, der Gewichtung des Erkenntnisinteresses, des Spielraumes qualitativer Methoden, der Verfolgung verschiedener Evaluationsziele, des Umganges mit Werten, der Beziehung zwischen Evaluatoren und Evaluierten, der Gewichtung von Generalisierbarkeit und Einzelfallanalyse, der Rolle von Fremd- und Selbstevaluation, der Nähe zu wissenschaftlicher Forschungsmethodik, und sie unterscheiden sich hinsichtlich der Partizipationsmöglichkeiten der Beteiligten. Zwischen einzelnen Varianten bestehen Ähnlichkeiten, zum Teil gehören sie aber verschiedenen Paradigmen mit gegensätzlichen deskriptiven und normativen Annahmen zum Gegenstandsbereich, zur Methodik, zu den Erkenntniszielen und -interessen an (Rindermann 2001; Beywl 1988).

Allen Varianten ist gemeinsam, dass über einen Gegenstandsbereich der Praxis durch methodisches Vorgehen relevante Informationen für die Gestaltung dieser Praxis ermittelt werden sollen. Wissenschaftliche Methodik *und* Praxisorientierung kennzeichnen die Evaluationsforschung (Kromrey 2003). Evaluation versteht sich somit als eine Form *rationaler Problemlösestrategie*, die für Entscheidungen zuverlässige empirische Datengrundlagen liefert sowie Entscheidungen mittels Klärung von Zielen und Dekomposition von komplexen Problemen rationaler gestalten möchte. Orientierung für Evaluation und Evaluationsforschung bieten neben gängiger sozialwissenschaftlicher Methodik (wie z.B. Versuchsplanung) eigens für den stärker praxis- und nutzenbezogenen Aspekt der Evaluation entwickelte Standards (Sanders/Beywl 2000).

2 Lehrevaluation, Lehrveranstaltungsevaluation und Lehrqualität

Anfang der 1990er Jahre begannen Universitäten im deutschen Sprachraum als Reaktion auf verschiedene Krisensymptome in der Hochschulausbildung und im Gefolge studentischer Proteste und Verwaltungsreformen der Länder die Qualität von Lehre und Ausbildung zu evaluieren. Als Verfahren etablierten sich neben *interner Evaluation* („Selbstevaluation“, „Lehrberichte“) und *externer Evaluation* durch Experten („niederländisches Modell“) vor allem *studentische Lehrveranstaltungsevaluationen*. Obwohl die Ergebnisse im Schnitt positiv ausfallen (Daniel 1996), bleibt Lehrevaluation umstritten, zudem trägt ihre Durchführung nicht immer zu einer Verbesserung der Lehre bei. Die Festlegung dessen, was unter der Qualität der Lehre zu verstehen ist, die Definition einzelner Merkmale und die Messung von Indikatoren dieser Qualität stellen ähnlich wie bei der Bestimmung der Qualität von Forschung ein schwieriges Unterfangen dar. Grundsätzlich sollten Lehr- oder Ausbildungsqualität als Ganzes und Veranstaltungsqualität als Teil dieser Gesamtqualität unterschieden werden. Beide Bereiche können z.B. mit Studierenden, Lehrenden (selbst) oder Fremdbeurteilern (geschulte Gutachter, Kollegen, Praktiker, Hochschuldidaktiker) als Einschätzern bewertet werden. Der Begriff *Lehrevaluation* ist veranstaltungs- und dozentenübergreifend und bezieht sich auf die Lehre insgesamt (inklusive Veranstaltungen, Ausbildungsinhalte, Bibliotheksausstattung, Studienzeiten, Absolventenchancen u.ä.), während unter *Lehrveranstaltungsevaluation* ausschließlich die Evaluation von Vorlesungen und Seminaren verstanden wird. Die Lehrveranstaltung stellt den Kern der Hochschullehre dar. In ihr findet der Wissens- und Kompetenzerwerb in Interaktion zwischen Dozenten und Studenten statt. In der Regel werden aber für beide Bereiche die umfassenderen und kürzeren Begriffe „Lehre“ und „Lehrevaluation“ verwendet. Der Begriff Evaluation impliziert, dass Bewertungen über reine Akzeptanzmaße inhaltlich hinausgehen sollen, dass die *Qualität* der Lehre, festgemacht an unterrichtswissenschaftlich überprüften Kriterien wie Struktur, studentischer Mitarbeit, einzelnen Rahmenbedingungen und Lerngewinn, bestimmt wird.

Abb. 1: Multidimensionales Bedingungsmodell des Lehrerfolgs (Rindermann 2001)



Zu den für erfolgreiche Lehrveranstaltungen verantwortlichen Lehrmethoden zählen als Ergebnis verschiedener Studien (direkte Befragungen, Regressionsanalysen mit Lehrevaluationsdaten, Faktorenanalysen, Strukturgleichungsmodelle, Auswertung von Konzepten guter Lehre, Unterrichtsforschung; Rindermann 2001, 1999) zuallererst Strukturierung und Klarheit. Als Determinanten auf Seiten des *Dozentenverhaltens* gelten darüber hinaus auch verständliche Erklärungen und Darstellungen, rhetorische Kompetenz, Motivierung, Verarbeitungstiefe, Zeit- und Schwierigkeitsmanagement, Engagement, Kooperativität/Kommunikation und in Seminaren das Leiten von Diskussionen. Weitere Bedingungsvariablen des Erfolges einer Veranstaltung stellen aber auch *studentische Determinanten* wie Fleiß, Vorwissen, die veranstaltungsinterne Beteiligung und Referate dar (Gold/Souvignier 2001). Dazu können *Kontextvariablen* (das gestellte Thema, die Anforderungen, Besuchszahl, Existenz einer Prüfung, Pflichtmäßigkeit des Besuches, Veranstaltungstyp u.ä. Rahmenbedingungen) den *Lehrerfolg* wie auch die Veranstaltungskritik beeinflussen (s. Abb. 1). Hinzu kommt auf Dozentenseite die wissenschaftliche Fachkompetenz, die durch Studierende aber nicht eingeschätzt werden sollte. Die Determinanten sind nicht unabhängig voneinander, der Dozent kann etwa Einfluss auf Thema, Anforderungen und Besuchsgrund ausüben, der Veranstaltungstyp auf Beteiligung und Betreuung, das Vorwissen auf Anforderungen, der Fleiß auf die Betreuung usw. Lehrerfolg wird multikriterial aufgefasst, innerhalb des Lehrerfolgs lässt sich zwischen Erfolgsmerkmalen für die Veranstaltung und intendierten Veränderungen auf Seiten der Studierenden unterscheiden. Das vier Komponenten unterscheidende Modell – drei

Bedingungsfaktoren und der Lehrerfolg – kann neben der Fragebogenkonstruktion auch als präskriptives Lehrmodell für die Beratung von Lehrenden herangezogen werden.

Eine differenzierte Beschreibung anhand theoretisch begründeter und empirisch überprüfter Skalen, die das komplexe Geschehen in Lehrveranstaltungen wiedergeben, ist Voraussetzung für Bewertung und sich hieran anknüpfende Maßnahmen.

3 Beziehung zwischen Forschung und Praxis in der Lehrevaluation

Es gibt neben der Psychotherapieevaluation kaum einen Bereich der Evaluationsforschung, der so eingehend untersucht wurde wie den der Lehr(veranstaltungs)evaluation. Desto mehr fällt ein dreifaches Defizit an Forschungsrezeption auf: Bei der Bewertung von Evaluationen und der Bewertung der Urteilsfähigkeit der Studierenden durch Hochschullehrer wird der Forschungsstand nicht ausreichend zur Kenntnis genommen (1), von Evaluationsforschern und Praktikern wird Hochschulforschung nur selektiv rezipiert (2) und bei der Planung und Durchführung von Lehrevaluationsprojekten durch Praktiker fließen Erfahrungen und Arbeiten der Wissenschaft kaum ein (3). Es liegt ein Praktizismus vor, ohne dass dieser der Praxis dient (Rindermann 2002).

3.1 Nichtbeachtung von Forschung bei der Bewertung von Lehrevaluation

In der Hochschulforschung wird die Lebenswelt der Wissenschaftler untersucht. Zwischen Praxis und Forschung besteht wie in der Hochschulforschung nur eine vage Trennung (Teichler 1994). Distanzierung von Alltagserfahrungen und Handlungsentlastetheit stellen aber Bedingungen wissenschaftlicher Rationalität dar. Die Evidenz lebensweltlicher Erfahrungen und der fehlende Bruch mit dieser Primärerfahrung legen dagegen eine perspektivische Interpretation aus Sicht eines Hochschullehrers nahe. Bei der Bewertung von Lehrevaluation und der Urteilsfähigkeit der Studierenden wird vornehmlich der Zugang über die eigene Erfahrung gewählt, nicht über die Wissenschaft. Warum ist dies im Falle der Lehrevaluation bedeutsam? Die Beurteilung von Hochschullehrern durch Studierende dreht die gängige Bewertungsrichtung um, sie führt zu einer (partiellen) Hierarchie-Inversion und verletzt die traditionelle Struktur einer historisch gewachsenen Institution. Die Einführung eines Lehrevaluations-Systems und die mögliche Berücksichtigung der Lehrqualifikation durch Evaluation bei der Stellenvergabe und Gehaltsbemessung würden die bisherigen Reproduktions- und Kooptationsmechanismen in der Wissenschaft und deren Autonomie gefährden. Mit erfolgreicher Integration in ein soziales System werden aber dessen Regeln und Strukturen stärker verteidigt. Erfolgreiche Sozialisation lässt sich durch ein hohes Maß an Symmetrie von objektiver mit subjektiver

Wirklichkeit kennzeichnen. Bei der Untersuchung von Einstellungen zur studentischen Lehrevaluation zeigte es sich dementsprechend, dass insbesondere Professoren skeptisch gegenüber der studentischen Urteilsfähigkeit eingestellt sind (Rindermann 2001).

3.2 Fachspezifische Rezeption von Forschung

Lehrevaluation ist Gegenstand der Forschung unterschiedlicher Disziplinen, und Praktiker entstammen unterschiedlichen Fachrichtungen (Helmke/Krapp 1999). Neben Psychologie, Soziologie, Pädagogik und Verwaltungswissenschaften sind auch innerhalb der genannten Fächer unterschiedliche inhaltliche und methodische Orientierungen erkennbar: an Pädagogischer Psychologie, Unterrichtsforschung, Diagnostik, Methodenlehre, Evaluationsforschung, Hochschulsozialisation, Hochschulforschung, Bildungsforschung usw. Ansätze der eigenen Arbeit und Publikationsorgan dieser, Literatursuche und Rezeption, Verständnis und Wertschätzung der Arbeiten anderer sind abhängig von Mechanismen fachlicher Binnensozialisation und Binnenselektion. Hochschulforschung ist wenig institutionalisiert und im deutschen Sprachraum fehlen Studiengänge wie „Higher Education“. Das wissenschaftliche Feld ist zersplittert, Evaluationsforscher und -praktiker orientieren sich weitgehend an Traditionen ihres eigenen Faches. So wird etwa in der Soziologie und Bildungsforschung kaum die stark empirisch-statistisch orientierte angloamerikanische Lehrevaluationsforschung zur Kenntnis genommen, in der Psychologie wiederum tauchen die stärker pädagogisch, soziologisch, organisations-, bildungs- und verwaltungswissenschaftlich geprägten Arbeiten der Hochschuldidaktiker und Hochschulforscher kaum auf. Dies obwohl verschiedene Forschungsfelder wie die der Lehrqualität oder der Hochschulsozialisation in Psychologie *und* Pädagogik *und* Soziologie fachlich angesiedelt sind (Helmke 2001).

3.3 Evaluationspraxis als formaler Verwaltungsakt

Fehlende Theorie-Praxis-Differenzierung und fachliche Zersplitterung machen aber nicht verständlich, dass bei der Durchführung von Evaluationsmaßnahmen der wissenschaftliche Kenntnisstand unabhängig von der Herkunft nur wenig umgesetzt wird. Bei Evaluationsmaßnahmen sind zunächst verschiedene Organisationsformen zu unterscheiden: als eigenständige Abteilungen oder Projekte, besetzt mit hauptamtlich für Evaluation zuständigen Akademikern in entweder ganzen oder halben Dauerstellen, in ganzen oder halben befristeten Projektstellen. Oder sie werden durch hauptamtliche Fachwissenschaftler aller an den Universitäten vertretenen Fächer durchgeführt, denen in einer zusätzlichen Nebentätigkeit in der eigenen Fakultät ähnlich der Aufgabe eines Institutsvorstandes und Dekans die Zuständigkeit für Lehrevaluation übertragen wurde oder die für Fachbereiche anderer Universitäten als externer kollegialer Gutachter wie bei der Manuskript- und Projektbewertung üblich angefragt wurden.

Gemeinsam ist diesen Varianten, dass Evaluation in der Praxis als der Universität vom Gesetzgeber oder Ministerium veranlasste Aufgabe verstanden wird, die gemäß der gegebenen Vorgaben zu erfüllen ist und deren Erfüllung nach innen, in die Universität, nach außen, in die eventuell interessierte Öffentlichkeit, und vor allem nach oben, in die die Auflagenerfüllung überprüfende Verwaltung, zu dokumentieren ist.

Ziel und Zweck einer Maßnahme – hier von Hochschulevaluation – können im Verlauf ihrer Durchführung verloren gehen. Evaluation als Verwaltungstätigkeit sollte möglichst störungsfrei verlaufen und im Sinne einer behördlichen Auflage erfolgreich abgearbeitet werden. In dieser Form wird die Evaluationspraxis der ihr eigenen Zielsetzung nicht gerecht: Praxis als tätige Auseinandersetzung mit der Wirklichkeit im Gegensatz zur Theorie ist immer ziel- und zweckbezogen, dient der Verfolgung eines subjektiven Sinns und objektiven Zwecks. Zwar ließe sich Evaluation als formaler Verwaltungsakt ohne Bewertungs-, Entscheidungs- und Verbesserungsperspektive immer noch ein Zweck unterstellen, sei es *institutionell* durch Erfüllung behördlicher Auflagen die Vermeidung von Problemen mit Ministerium und Verwaltungsinstanzen oder sei es *persönlich* die Möglichkeit, eine bezahlte Anstellung innezuhaben, sich durch Übernahme der Evaluationstätigkeit anderer Aufgaben zu entledigen, oder sei es *wissenschaftlich* die Datengewinnung für die Forschung. Der Zweck der Evaluation, durch *Informationsgewinnung* mittels systematischer Analyse und empirischer Untersuchung von Konzepten, Bedingungen, Prozessen und Wirkungen zielgerichteter Aktivitäten zu deren *Bewertung* und vor allem zu deren *Verbesserung* beizutragen (Hager et al. 2000; Rindermann 2001), geht aber verloren. Evaluation als Verwaltungsaufgabe ohne Qualitätsorientierung könnte genau das Gegenteil dessen erreichen, was sie (im engeren Sinne) bezwecken soll: Statt Verbesserung der Lehre und der Ausbildung deren Verschlechterung durch vermehrte Arbeitsbelastung für Wissenschaftler und Lehrende außerhalb der universitären Kernbereiche von Forschung und Lehre.

4 Wichtige Ergebnisse der bisherigen Forschung: Urteilskompetenz der Studierenden, Objektivität, Reliabilität, Generalisierbarkeit und Validität

Von Skeptikern studentischer Lehrevaluationen wird oft die Urteilskompetenz der Studierenden in Frage gestellt. Konkret wird die Heterogenität der Urteile oder mangelnde Vergleichbarkeit von Veranstaltungskritik aus Veranstaltungen mit verschiedenen Themen und vor allem der Einfluss durch Drittvariablen wie Thema, Besuchsgrund und Benotungsstrenge betont. Es wird befürchtet, Studierende würden nur nach Beliebtheit des Dozenten oder des Themas eindimensional urteilen.

Ergebnisse verschiedener Untersuchungsansätze (Interraterreliabilität; Reliabilitätsberechnungen; Faktorenanalysen; Generalisierbarkeit; Korrelation verschiedene Aspekte der Lehre abdeckender Skalen mit potentiellen Verzerrungsfaktoren; Ver-

gleich der Resultate mit den Urteilen geschulter Fremdurteiler; Zusammenhang mit Leistungstestergebnissen) sprechen für die Brauchbarkeit von Lehrevaluationen. Die Reliabilitäten für Kursmittel sind zufriedenstellend bis gut (r um .80). In Faktorenanalysen und Generalisierbarkeitsüberprüfungen (dimensionsbezogene Vergleiche der Urteilsstabilität von Evaluationsergebnissen, die aus verschiedenen Veranstaltungen eines Dozenten stammen, mit Stabilitäten der Beurteilungen für Veranstaltungen gleichen Themas, die verschiedene Dozenten hielten) lassen sich das Handeln des Dozenten (z.B. Lehrkompetenz), Lehrerfolgsscharakteristika (z.B. Allgemeinbeurteilung der Veranstaltung und Lernen), studentische Beteiligung und Rahmenbedingungsvariablen (z.B. Anforderungen) unterscheiden. Voraussetzung hierfür ist ein multidimensionales Instrument.

Geben Studierende in ihren Urteilen tatsächlich das Veranstaltungsgeschehen wieder (Validität)? Zwischen Fremdurteilen bzw. Leistungsmaßen als Kriterium und studentischen Urteilen konnten Zusammenhänge in einer Höhe wie zwischen Intelligenztests und Leistungskriterien bekannt festgestellt werden. Beurteilungen durch Studierende zeichnen somit kein unrealistisches Bild der Lehre und sind sensibel für Lerngewinn. Zudem reflektieren Dozenten-Skalen mehr das Lehrverhalten des Lehrenden als Rahmenbedingungen. Deshalb kann von studentischen Lehrevaluationen als einem Maß universitärer *Lehrqualität* gesprochen werden.

Auch wenn die Frage der Kausalität nicht abschließend geklärt ist, kann ein gewisser Einfluss von Veranstaltungsthema und Besuchsgrund auf die Veranstaltungsbeurteilung und – zwar geringer, aber inhaltlich gravierender – auf die Dozentenbeurteilung nicht ausgeschlossen werden. Generell ist deshalb für die Praxis zu empfehlen, Evaluationen in mehreren verschiedenen Veranstaltungen (Mindeststichprobengröße jeweils $N=10-15$ Studierende aus fünf Veranstaltungen) vorzunehmen, falls es Ziel ist, die Lehrqualität eines Dozenten zu bestimmen. Der Einfluss eventuell auftretender Störvariablen kann so minimiert, aber nicht völlig ausgeschlossen werden.

5 Nutzen und Qualitätsentwicklung durch Lehrevaluation

Studentische Veranstaltungskritik sollte einen Beitrag zur Verbesserung der Lehre leisten, Evaluation ist nicht Selbstzweck. Kann Lehre durch Evaluationsmaßnahmen nachweisbar verändert werden? Stellt Evaluation als solche schon eine Intervention dar (Henninger 2000)? – Die Summe bisheriger Studien zeigt, dass bloße Durchführung von Lehrevaluation nicht ausreichend ist für eine deutliche Verbesserung der Lehre (Rindermann/Kohler 2003). Alles andere wäre auch unwahrscheinlich und würde Ergebnissen sozialwissenschaftlicher Forschung in anderen Bereichen widersprechen: Umweltbedingungen blieben stabil, Routinen sind nur durch Information nicht aufbrechbar, Evaluation stellt zunächst Messung dar, noch keine Intervention!

Ergebnisse verweisen auf die Notwendigkeit, Feedback studentischer Lehrevaluationen mit mündlicher Erläuterung und persönlicher Beratung zu ergänzen.

Dann sind positive wie substantielle Veränderungen im Messwiederholungsdesign oder verbesserte Einschätzungen im Vergleich zu Kontrollgruppen erkennbar. Diese treten vor allem in solchen Bereichen auf, die durch Handeln des Dozenten direkt veränderbar sind (für Heidelberger Inventar zur Lehrveranstaltungsevaluation, HILVE: Lehrkompetenz: $d=0,94$; Dozentenengagement: $d=0,78$) und in Skalen, die den Erfolg der Lehre aus Sicht der Studierenden wiedergeben (für HILVE: Allgemeinurteil: $d=0,85$; Rindermann/Kohler 2003). Evaluation dient in solchen Ansätzen der Bestimmung von Defiziten (wo besteht Handlungsbedarf?), sie dient als Handlungsanstoß und der Motivationsänderung (Balk 2000) und sie dient als Informationsgrundlage für Beratung und Training. Hierfür sind Normwerte unabdingbar. Zusätzlich prüft sie den Erfolg von Rückmeldung, Beratung und Training. Ohne Anreize für gute Lehre bleibt Engagement für die Lehre aber wahrscheinlich auf individuelle Voraussetzungen (Interesse u.ä.) beschränkt (Leitner 1998; Spiel/Fischer 1998).

Evaluationsprojekte, die Lehre und Studium ganzer Fachbereiche untersuchen, haben bislang nicht in einer gängigen wissenschaftlichen Kriterien genügenden Form Veränderungserfolge belegen können. Generell ist eine wissenschaftliche Fundierung und Überprüfung bei dieser Praxis der Evaluation seltener: „Irgendwelche wissenschaftlichen Methoden zur Erhebung ... oder Auswertung der herangezogenen Daten und Statistiken werden weder gefordert noch gefördert.“ (Westermann 2002, S. 20) Ähnlich kritisch äußerte sich Weinert (2001, S. 6): „Stark gefördert von einigen Bundesländern und nachhaltig unterstützt durch Initiativen der Hochschulrektorenkonferenz entsteht ... eine Evaluationspraxis, die relativ unabhängig vom jeweiligen Stand der theoretischen Reflexion ist.“

6 Anwendungsvorschläge

Lehrevaluation und Lehrveranstaltungsevaluation sollten in ein *allgemeines Verfahren zur Bestimmung und Förderung* von Lehr-, Ausbildungs- und Forschungsqualität integriert werden. Die alleinige Einschätzung von Lehrqualität ist nicht ausreichend für eine Verbesserung. Insbesondere ist aus Effektivitätsgründen wie auch aus ethischen Gründen ein Ansatz abzulehnen, Evaluation – die bloße Messung der Lehrqualität – ohne Möglichkeit der Qualitätsverbesserung durchzuführen. Eine Institution zu evaluieren macht nur dann Sinn, wenn Möglichkeiten bestehen, auf Ergebnisse zu reagieren und förderliche Veränderungen einzuleiten. Qualität könnte sonst sogar vermindert werden, weil Zeit für Forschung und Lehre verloren geht. So sollte Lehrevaluation außer in problematischen Fällen oder bei Interventionsmaßnahmen nicht öfters als alle zwei bis vier Jahre je Dozent und sechs bis zehn Jahre je Fachbereich durchgeführt werden.

Ist man an der Verbesserung der Qualität der Lehre in Form von Lehrveranstaltungen interessiert, sollte die Lehre im Rahmen eines *Beratungs-Rückmelde-Verfahrens* evaluiert werden. Die private Nutzung von Lehrveranstaltungsevaluation kann der Selbstvergewisserung über den eigenen Hochschulunterricht dienen, in

günstigen Fällen lässt sich mittels dadurch initiiertes Kommunikationsprozesse auch ohne Beratung die Lehre verbessern. In der Regel gehen eher die motivierteren Dozenten und vor allem diejenigen Dozenten diesen Weg, die gute bzw. gut beurteilte Lehre halten. Ist Verbesserung der lehrveranstaltungübergreifenden Ausbildungsqualität das Ziel, sind interne Selbst- und externe Fremdevaluationsprozesse durch Kollegen (peers) sowie studentische Einschätzungen in *effektive Kommunikations- und Entscheidungsprozesse* zu überführen, an deren Ende jeweils Anreize für gute Lehre stehen müssen. Moderne Instruktions-, Management- und Organisationskonzepte sowie Ergebnisse der Hochschul- und Lehrevaluationsforschung liegen hierfür bereit (z.B. Theall/Franklin 1991; Webler 2002; Wottawa 2001). Eine Unterstützung durch externe Experten der Hochschul- und Evaluationsforschung ist anzuraten. Die Effektivität eines solchen Verfahrens für Lehrqualität muss über Erfahrungsberichte hinausgehend in wissenschaftlicher Form geprüft werden.

Beispiel: Rückmelde- und Beratungsverf. in Nürnberg durch Rindermann/Dresel

1. Rückmeldung gruppenstatistischer Ergebnisse in einer Versammlung der Dozenten (hier: außerhalb des täglichen Lehrbetriebes in einer externen Tagungsstätte mit Verpflegung; Dauer inkl. individuellem Beratungsgespräch insgesamt ca. sechs Stunden)
2. Übermittlung des individuellen Feedbacks durch Berater:
 - Feedbackpräsentation und -erläuterung
 - Erfragung der Sichtweise des Dozenten
 - Ist-Analyse, Abklärung des Istzustandes
 - Identifizierung und Herausstreichung von Stärken, Stützung
 - Identifizierung von kritischen Aspekten der Lehre
 - gemeinsame Definition von Verbesserungszielen
 - Möglichkeiten zur Verbesserung erarbeiten, Vorschläge formulieren
 - konkrete Schritte vereinbaren
 - Zusammenfassung des Gesprächs
 - Motivierung Veränderungsschritte einzuleiten, Ermutigung, Hinweis auf wiederholte Evaluation zu einem späteren Zeitpunkt
3. erneute Lehrevaluation zur Bestimmung von Veränderungen

Nicht übersehen werden sollte hierbei jedoch, dass Universitäten als Institution und insbesondere ihr Auftrag der Erkenntnisgewinnung und Ausbildung wie Bildung nur eingeschränkt effektivitäts- und effizienzorientiert steuerbar sind: Die Grenzen eines an Effektivität, Effizienz und Meritokratie orientierten Steuerungssystems liegen an der inhaltlich schwierigen *Messbarkeit* von Lehr-, Forschungs- und Hochschulqualität – quantitative Indikatoren sagen etwas über Qualität aus, bilden diese aber nicht vollständig ab, quantitative Indikatoren und Qualitätsurteile können durch Störfaktoren beeinflusst werden, lassen sich bei Verbindung mit Gratifikationen einseitig maximieren (z.B. bei Publikationszählung durch Publikationszersplitterung Erhöhung der Publikationsrate), sie sind abhängig von Beziehungsmustern zwischen Urteilernden und Beurteilten, deren wissenschaftlichen Orientierungen usw. Studentische Lehrveranstaltungsevaluation ist hierbei vergleichsweise aussagekräftig und diese ist gut erforscht. Die Grenzen liegen neben Bestimmungsproblemen aber auch in *Umset-*

zungsdefiziten und im Beharrungsvermögen einer behördlich organisierten Institution, in *ethischen Problemen* (können Personen, Fächer oder Universitäten mit unterschiedlichen Eingangsvoraussetzungen einheitlich hinsichtlich ihrer Leistung bewertet werden?) und in Grenzen, die durch eine *ingeschränkte Funktionalität des meritokratischen Modells* für wissenschaftliche Arbeit und Universitäten bedingt sind: Forschende und unterrichtende Tätigkeit, die nicht im Wesentlichen selbstmotiviert und selbstgesteuert durchgeführt werden, würden als nur extern gesteuerte zur langfristig unproduktiven Auftragsarbeit. Die universitäre Freiheit vor äußeren Eingriffen sichert einen gewissen Raum vor externen staatlichen, gesellschaftlichen und kollegialen Eingriffen durch Ministerien, Moden, externe Geldgeber, Wissenschaftsverbände und Kollegien, der sowohl privatisierenden Missbrauch ermöglicht als auch unkonventionellen wissenschaftlichen Wegen ihre Freiheit lässt. Vielfalt kann entstehen, neue Ideen entwickeln sich. Allerdings ist unverkennbar, dass ein Hochschulsystem, das gänzlich auf die Mittel der Steuerung durch Evaluation und Qualitätsmanagement, durch Qualifizierung und Gratifikation verzichtet, sich damit auch wirksamer Möglichkeiten enthält, Qualität in Forschung und Lehre zu erhöhen.

Theorie und Praxis, Forschung und Anwendung sind stärker wechselseitig zu verknüpfen. Forschung sollte neben Lehrveranstaltungsevaluationen auch die Effektivität von Evaluationen ganzer Fachbereiche, die Auswirkungen von Evaluation auf Qualität und Entwicklung dieser Qualität untersuchen. Die Praxis wiederum sollte Ergebnisse dieser Forschung für die Gestaltung dieser Praxis berücksichtigen und nach Möglichkeit selbst ihr Handeln einer wissenschaftlichen Analyse zum Nutzen von Universität und Evaluation unterziehen.

Literatur

- Balk, Michael, 2000: Die Evaluation von Lehrveranstaltungen. Die Wirkung von Evaluationsrückmeldung. Frankfurt a.M.: Peter Lang.
- Beywl, Wolfgang, 1988: Zur Weiterentwicklung der Evaluationsmethodologie: Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation. Frankfurt: Peter Lang.
- Daniel, Hans-Dieter, 1996: Evaluierung der universitären Lehre durch Studenten und Absolventen. In: Zeitschrift für Sozialisationsforschung und Erziehungssoziologie, 16(2), S. 149-164.
- Gold, Andreas; Souvignier, Elmar, 2001: Referate in Seminaren: Warum man sie beibehalten und verbessern sollte. Das Hochschulwesen, 3, S. 70-74.
- Hager, Willi; Patry, Jean-Luc; Brezing, Hermann (Hrsg.), 2000: Evaluation psychologischer Interventionsmaßnahmen. Bern: Huber.
- Helmke, Andreas, 2001: Hochschulsozialisation. In: Detlef Rost (Hrsg.): Handwörterbuch Pädagogische Psychologie. Weinheim: PVU, S. 255-260.
- Helmke, Andreas; Krapp, Andreas, 1999: Lehren und Lernen in der Hochschule. In: Zeitschrift für Pädagogik, 45(1), S. 19-24.

- Henninger, Michael, 2000: Evaluation: Diagnose oder Therapie. In: C. Harteis; H. Heid; S. Kraft (Hrsg.): Kompendium Weiterbildung. Aspekte und Perspektiven betrieblicher Personal- und Organisationsentwicklung. Opladen: Leske & Budrich, S. 249-260.
- Kromrey, Helmut, 2003: Evaluierung und Evaluationsforschung: Begriffe, Modelle und Methoden. In: Psychologie in Erziehung und Unterricht, 50, S. 11-26.
- Leitner, Erich, 1998: The pedagogical qualification of the academic teaching staff and the quality of teaching and learning. In: Higher Education in Europe, 23(3), S. 339-349.
- Marsh, Herbert, 1987: Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. In: International Journal of Educational Research, 11, S. 253-388.
- Rindermann, Heiner, 1999: Was zeichnet gute Lehre aus? Ergebnisse einer offenen Befragung von Studierenden und Lehrenden nach Merkmalen guter Dozenten und Veranstaltungen. In: Zeitschrift für Hochschuldidaktik, 23(1), S. 136-156.
- Rindermann, Heiner, 2001: Lehrevaluation – Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen. Mit einem Beitrag zur Evaluation computerbasierten Unterrichts. Landau: Empirische Pädagogik.
- Rindermann, Heiner, 2002: Lehrevaluation an Hochschulen: Schlussfolgerungen aus Forschung und Anwendung für Hochschulunterricht und seine Evaluation. Unveröffentlichtes Manuskript.
- Rindermann, Heiner; Kohler, Jürgen, 2003: Lässt sich die Lehrqualität durch Evaluation und Beratung verbessern? Überprüfung eines Evaluations-Beratungs-Modells. In: Psychologie in Erziehung und Unterricht, 50(1), S. 71-85.
- Rossi, Peter Henry; Freeman, Howard; Lipsey, Mark, 1999: Evaluation – a systematic approach, 6th edition. Newbury Park: Sage.
- Sanders, James; Beywl, Wolfgang (Hrsg.), 2000: Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“. Opladen: Leske + Budrich.
- Spiel, Christiane; Fischer, Ursula, 1998: Evaluierung eines Weiterbildungsangebots für Hochschullehrende. In: Zeitschrift für Hochschuldidaktik, 22(1), S. 83-99.
- Teichler, Ulrich, 1994: Hochschulforschung – Situation und Perspektiven. Das Hochschulwesen, 4, S. 169-177.
- Theall, Michael; Franklin, Jennifer (Hrsg.), 1991: Effective practices for improving teaching. San Francisco: Jossey-Bass (New directions for teaching and learning 48).
- Webler, Wolff-Dietrich, 2002: Vergleich von fünf wichtigen Evaluationsmodellen in Deutschland – Grundlagen einer Meta-Evaluation. In: Karl Neumann; Jürgen Osterloh (Hrsg.): Gute Lehre in der Vielfalt der Disziplinen. Weinheim: Beltz, S. 141-172.
- Weinert, Franz, 2001: Die evaluierte Universität. Heidelberg: Universitätsrede am 25. Januar 2001.
- Westermann, Rainer, 2002: Merkmale und Varianten von Evaluationen: Überblick und Klassifikation. In: Zeitschrift für Psychologie, 210(1), S. 4-26.
- Wottawa, Heinrich, 2001: Qualitätsmanagement durch Zielvereinbarung. In: Christiane Spiel (Hrsg.): Evaluation universitärer Lehre. Münster: Waxmann, S. 151-164.
- Wottawa, Heinrich; Thierau, Heike, 1998: Lehrbuch Evaluation. Bern: Hans Huber.

Dr. Heiner Rindermann
Institut für Psychologie
Otto-von-Guericke-Universität Magdeburg
Postfach 4120
39016 Magdeburg
Tel.: ++49.391-6711919
Fax.: ++49.391-6711965
eMail: heiner.rindermann@gse-w.uni-magdeburg.de

Dr. Heiner Rindermann (Jahrgang 1966) studierte Psychologie in Heidelberg und schloss dieses Studium 1993 mit Diplom und 1995 mit Promotion ab. 1994-1999 arbeitete er an der Universität München im Institut für Pädagogische Psychologie in einem Forschungsprojekt zum Thema Hochbegabung. Seit 1999 lehrt und forscht er als Wissenschaftlicher Assistent im Bereich Methodenlehre, Diagnostik und Evaluation am Psychologischen Institut der Universität Magdeburg; 2001 Gastprofessur in Graz. Er beschäftigt sich mit den Themen Lehrevaluation und Unterrichtsforschung, Intelligenz und Hochbegabung, Selbstkonzept und Persönlichkeit, evolutionäre Psychologie und Wissenschaftsforschung. Aktuelle Veröffentlichung: Rindermann, Heiner; Kohler, Jürgen, 2003: Lässt sich die Lehrqualität durch Evaluation und Beratung verbessern? Überprüfung eines Evaluations-Beratungs-Modells. In: Psychologie in Erziehung und Unterricht, 50(1), S. 71-85.