

Qualität in der Wissenschaft

Zeitschrift für Qualitätsentwicklung in
Forschung, Studium und Administration

- Der Journal Impact Factor
– Ein problematischer bibliometrischer Indikator
- Modellierung von Qualitätsaspekten
in Hochschulprozessen
- Pilotprojekt „Lehrveranstaltungsbewertung mittels
Beobachtungsverfahren an der MedUni Wien“
- Formative und summative Evaluation
schärfer unterscheiden!
Anmerkungen über den Unterschied zwischen
Rückmeldung und (summativer) Evaluation

2
2012

Herausgeberkreis

Hans-Dieter Daniel, Dr., Professor für Sozialpsychologie und Hochschulforschung, ETH Zürich (CH), Leiter der Evaluationsstelle der Universität Zürich

Michael Heger, Dr., Evaluationsbeauftragter der Fachhochschule Aachen, Leiter des Bereichs Hochschuldidaktik und Evaluation in der zentralen Qualitätsentwicklung ZQE

Achim Hopbach, Dr., Geschäftsführer der Stiftung zur Akkreditierung von Studiengängen in Deutschland, Bonn

Stefan Hornbostel, Dr., Leiter des Instituts für Forschungsinformation und Qualitätssicherung (iFQ), Berlin, Professor für Soziologie (Wissenschaftsforschung), Institut für Sozialwissenschaften der Humboldt-Universität zu Berlin

René Krempkow, Dr., Institut für Forschungsinformation und Qualitätssicherung (iFQ), Berlin

Sandra Mittag, Dr., Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung, München

Philipp Pohlenz, Dr., Geschäftsführer des Zentrums für Qualitätsentwicklung in Lehre und Studium an der Universität Potsdam

Uwe Schmidt, Dr., Leiter des Zentrums für Qualitätssicherung und -entwicklung der Universität Mainz, Geschäftsführer des Hochschulevaluationsverbundes Südwest

Wolff-Dietrich Webler, Dr., Leiter des Instituts für Wissenschafts- und Bildungsforschung Bielefeld (IWBB), Professor of Higher Education, University of Bergen (Norway); Staatliche Pädagogische Universität Jaroslavl/Wolga

Don Westerheijden, Dr., Center for Higher Education Policy Studies (CHEPS), University of Twente, Netherlands

Lothar Zechlin, Dr., ehem. Rektor der Universität Duisburg-Essen, Professor für Öffentliches Recht und Politikwissenschaft, Vertreter der Agenturen im Akkreditierungsrat

Hinweise für die Autoren

In dieser Zeitschrift werden i.d.R. nur Originalbeiträge publiziert. Sie werden doppelt begutachtet. Die Autor/innen versichern, den Beitrag nicht zu gleicher Zeit an anderer Stelle zur Publikation angeboten zu haben. Beiträge werden nur dann angenommen, wenn die Autor/innen den Gegenstand nicht in vergleichbarer Weise in einem anderen Medium behandeln. Senden Sie bitte das Manuskript als Word-Datei und Abbildungen als JPG-Dateien per E-Mail an die Redaktion (Adresse siehe Impressum).

Wichtige Vorgaben zu Textformatierungen und beigefügten Fotos, Zeichnungen sowie Abbildungen erhalten Sie in den „Autorenhinweisen“ auf unserer Verlags-Homepage: „www.universitaetsverlagwebler.de“.

Ausführliche Informationen zu den in diesem Heft aufgeführten Verlagsprodukten erhalten Sie ebenfalls auf der zuvor genannten Verlags-Homepage.

Impressum

Anschrift Verlag, Redaktion, Abonnementenverwaltung:

UVW UniversitätsVerlagWebler
Der Fachverlag für Hochschulthemen
Bünder Straße 1-3 (Hofgebäude), 33613 Bielefeld
Tel.: 0521-92 36 10-12, Fax: 0521-92 36 10-22

Satz: UVW, info@universitaetsverlagwebler.de

Anzeigen: Die Zeitschrift „Qualität in der Wissenschaft“ veröffentlicht Verlagsanzeigen, Ausschreibungen und Stellenanzeigen. Aufträge sind an den Verlag zu richten. Die jeweils gültigen Anzeigenpreise sind der Homepage erhalten Sie auf Anfrage beim Verlag.

Redaktionsschluss dieser Ausgabe: 15.09.2012

Umschlaggestaltung: Wolff-Dietrich Webler, Bielefeld
Gesetzt in der Linotype Syntax Regular

Druck: Sievert Druck & Service GmbH,
Potsdamer Str. 190, 33719 Bielefeld

Abonnement/Bezugspreis:

Jahresabonnement: 68 Euro zzgl. Versandkosten
Einzelpreis: 17,25 Euro zzgl. Versandkosten

Abo-Bestellungen und die Bestellungen von Einzelheften sind unterschrieben per Post oder Fax bzw. per E-Mail an den Verlag zu richten. Eine Abo-Bestellvorlage finden Sie unter www.universitaetsverlagwebler.de.

Das Jahresabonnement verlängert sich automatisch um ein Jahr, wenn es nicht 6 Wochen vor Jahresende gekündigt wird.

Erscheinungsweise: 4mal jährlich

Copyright: UVW UniversitätsVerlagWebler

Die mit Verfassernamen gekennzeichneten Beiträge geben nicht in jedem Falle die Auffassung der Herausgeber bzw. Redaktion wieder. Für unverlangt eingesandte Manuskripte/Rezenzionsexemplare wird keine Verpflichtung zur Veröffentlichung/Besprechung übernommen. Sie können nur zurückgegeben werden, wenn dies ausdrücklich gewünscht wird und ausreichendes Rückporto beigefügt ist. Die Urheberrechte der hier veröffentlichten Artikel, Fotos und Anzeigen bleiben bei der Redaktion. Der Nachdruck ist nur mit schriftlicher Genehmigung des Verlages gestattet.

Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Qualität in der Wissenschaft

Zeitschrift für Qualitätsentwicklung in
Forschung, Studium und Administration

Einführung des geschäftsführenden Herausgebers

29

Wolff-Dietrich Webler

Formative und summative Evaluation schärfer
unterscheiden! Anmerkungen über den
Unterschied zwischen Rückmeldung und
(summativer) Evaluation

48

Qualitätsentwicklung/-politik

Werner Marx & Lutz Bornmann

Der Journal Impact Factor –
Ein problematischer bibliometrischer Indikator

30

Heinrich Seidlmeier

Modellierung von Qualitätsaspekten
in Hochschulprozessen

35

Forschung über Qualität in der Wissenschaft

Isolde Kirnbauer & Katharina Stowasser-Bloch

Pilotprojekt „Lehrveranstaltungsbewertung mittels
Beobachtungsverfahren an der MedUni Wien“

40

Meldungen

54

Seitenblick auf die Schwesterzeitschriften

Hauptbeiträge der aktuellen Hefte
Fo, HSW, HM, P-OE und ZBS

IV

im UniversitätsVerlagWebler erhältlich:

Sandra Mittag, Rüdiger Mutz & Hans-Dieter Daniel:

Institutionelle Qualitätssicherung der Lehre auf dem Prüfstand: Eine Fallstudie an der ETH Zürich

Reihe: Qualität - Evaluation - Akkreditierung

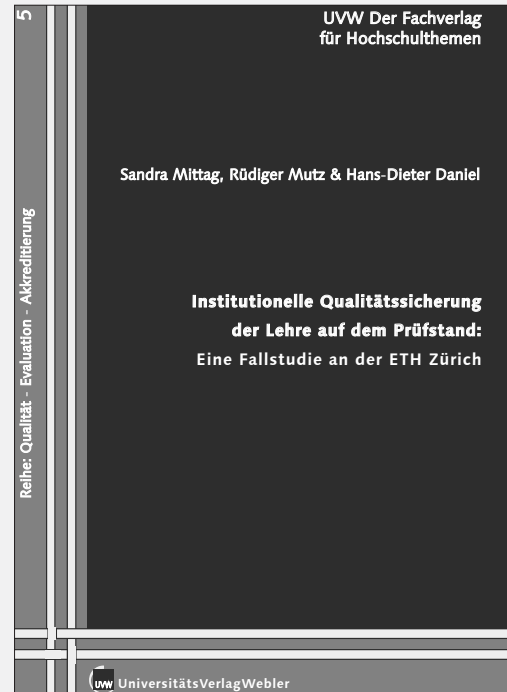
Im Rahmen der vorliegenden Studie wurde das Qualitätssicherungssystem der ETH Zürich im Bereich Lehre einer umfassenden Meta-Evaluation unterzogen.

Das Qualitätssicherungssystem stützt sich auf die vier Instrumente Lehrveranstaltungsbeurteilung, Absolventenbefragung, Selbstevaluation und Peer Review.

Die Ergebnisse zeigen unter anderem, dass die ETH Zürich über etablierte Qualitätssicherungsinstrumente verfügt, die weitestgehend akzeptiert sind.

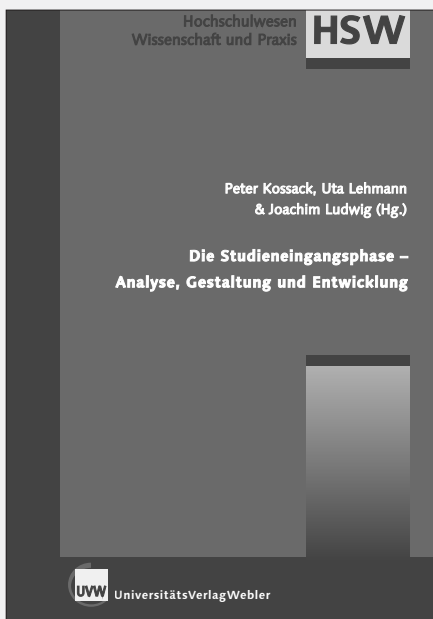
Allerdings bestehen bei allen vier Instrumenten Optimierungspotentiale.

*ISBN 3-937026-74-6, Bielefeld 2012, 115 S.,
19,50 Euro zzgl. Versandkosten*



Bestellung - E-Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

Peter Kossack, Uta Lehmann & Joachim Ludwig (Hg.): Die Studieneingangsphase – Analyse, Gestaltung und Entwicklung



Der vorliegende Band versammelt eine Reihe von Arbeiten, die im Kontext der Weiterentwicklung der Qualität von Lehre entstanden sind. Dabei wird im Besonderen die Studieneingangsphase als zentrale Übergangsstelle in Bildungsbiographien in den Blick genommen. Die Arbeiten reichen von der Vorstellung einer empirisch fundierten Analyse typischer Problemlagen in Studieneingangsphasen über die Darstellung von Instrumenten zur Entwicklung von Studieneingangsphasen bis hin zur kritischen Reflexion der Studieneingangsphasenpraxis.

Vor dem Hintergrund der Umstellung von Studiengängen im Zuge des Bologna-Prozesses geben die Beiträge Einblick zu aktuellen Anforderungen und Problemstellungen, mit denen Studiengangplanende, Hochschullehrende wie auch Studierende in der Studieneingangsphase konfrontiert sind. Darüber hinaus werden für eine Entwicklung von Studiengängen und die Gestaltung der Hochschullehre relevante Potentiale und Handlungsmöglichkeiten aufgezeigt.

ISBN 10: 3-937026-77-0, Bielefeld 2012, 165 Seiten, 19.80 Euro

Bestellung - E-Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

Reihe Hochschulwesen: Wissenschaft und Praxis

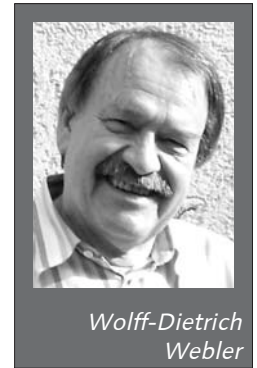
In zahlreichen Alltagssituationen in der Wissenschaft geht es heute um die Bewertung individueller Leistungen. Das kann in Berufungsverfahren bzw. Bleibeverhandlungen sein, aber auch bei der Neuaushandlung von individuellen Gehältern (wenn auch nicht im deutschen Wissenschaftssystem). Aber es kann sich auch um den einfachen Vorgang der Literatursuche und der Beurteilung der Relevanz von Publikationen handeln, die die Datenbanken breit stellen und die nicht alle einzeln auf ihre Qualität geprüft werden können. Die Unübersichtlichkeit der Fluten von Wissenschaftspublikationen weltweit hat ein Interesse daran geweckt, schon am Publikationsort – etwa der Zeitschrift – erkennen zu können, wie relevant ein Beitrag vermutlich sein wird oder umgekehrt vor allem diese Zeitschriften nach einschlägigen Artikeln für ein eigenes Vorhaben absuchen zu können. Dafür wird seit 1975 zunehmend der Journal Impact Factor (JIF) verwendet – auch wenn er ursprünglich zu anderen Zwecken entstanden ist. Dieser Faktor gewinnt großen Einfluss und lohnt daher eine nähere Betrachtung. *Werner Marx & Lutz Bornmann* rekonstruieren in ihrem Beitrag **Der Journal Impact Factor – Ein problematischer bibliometrischer Indikator** seine Entstehung und analysieren, ob die Wirkungen durch die Datengrundlagen gerechtfertigt sind. Das Ergebnis ist ernüchternd.

Seite 30

Die Durchführung von Berufungsverfahren (aber auch anderer Verfahren, wie der Einrichtung neuer Studiengänge, Akkreditierungen, Evaluierungen oder Prüfungen) ist unter Beachtung aller rechtlichen Vorgaben ein anspruchsvoller und umfangreicher, zahlreiche Stationen umfassender Prozess. Um seine Transparenz zu erhöhen und gleichzeitig qualitätssichernde Aspekte stärker berücksichtigen zu können, hat *Heinrich Seidlmeier* in seinem Beitrag **Modellierung von Qualitätsaspekten in Hochschulprozessen** Anleihen bei der Prozessmodellierung gemacht. Der Autor zeigt, „wie Modelle für Hochschulprozesse mit qualitätskritischen Aspekten angereichert werden können“ (Einleitung).

Seite 35

Die Qualität einer Lehrveranstaltung festzustellen ist der erste Schritt zu ihrer Sicherung bzw. Verbesserung. Aber worin besteht deren Qualität? Wenn die Lehrenden und Studierenden zu unterschiedlichen Bewertungen der Veranstaltung kommen, entwickeln Lehrende manchmal den Wunsch, von neutraler Seite die Qualität „objektiv“ feststellen zu lassen. Stabsstellen für Qualitätsmanagement können dann u.U. mit erheblichen Handlungserwartungen konfrontiert werden. Um einem solchen Wunsch nachzukommen, hat die Stabsstelle für Evaluation und Qualitätsmanagement an der Medizinischen Universität Wien ein Pilotprojekt durchgeführt. *Isolde Kirnbauer & Katharina Stowasser-Bloch* stellen in ihrem Beitrag **Pilotprojekt „Lehrveranstaltungsbewertung mittels Beobachtungsverfahren an der MedUni Wien“** die Entwicklung und Ergebnisse vor. Es ist positiv,



Wolff-Dietrich
Webler

dass ihr Ansatz, die methodischen Überlegungen und praktischen Maßnahmen eines solchen Vorhabens einmal durchgespielt worden sind und ihre Erfahrungen aus der Pilotphase berichtet werden. Trotzdem ist dieser Ansatz in seiner Grundanlage und seinem Gesamtaufwand diskussionsbedürftig. Um ihn eben diesem Diskurs zuzuführen und in der Fachgemeinschaft diskutierbar zu machen, wie das dem Konzept dieser Zeitschrift entspricht, wird der Bericht hier veröffentlicht. Sicherlich stellen sich ähnliche Fragen oder Ansinnen häufiger auch an anderen Hochschulen.

Seite 40

Daran schließt der nachfolgende Beitrag direkt an. Nicht das Projekt als solches wird diskutiert, aber einige der darin liegenden Grundsatz- und Methodenfragen werden zum Anlass genommen, sie gründlicher zu betrachten. Man kann diesem Wunsch nach einer „objektiveren“ Bewertung von Lehrveranstaltungen, wie in dem vorstehenden Beitrag erprobt, kritisch bis ablehnend gegenüber stehen. Möglichen Einwänden gegen den Ansatz, die Grundanlage und den Gesamtaufwand geht der nachfolgende Artikel nach. Oft wird in Evaluationsverfahren nicht deutlich getrennt zwischen dem Zweck, einen laufenden Lehr-/Lernprozess zu optimieren (also einer formativen Evaluation) einerseits und dem Versuch, zu einem bestimmten Zeitpunkt quasi bilanzierend die Qualität einer Lehrveranstaltung festzustellen andererseits (also einer summativen Evaluation). Methodisch bringt das erhebliche Differenzen mit sich. Die nächste Falle steckt im Focus der Evaluation. Außer bei Personalbeurteilungen (z.B. in Personalauswahl-, etwa Berufungsverfahren) steht nicht allein die Lehrperson bzw. ihre Lehre im Focus, sondern die Veranstaltung. Die dritte Falle wartet in den Bewertungsmaßstäben. Was bedeutet die Qualität einer Lehrveranstaltung? Für die Beantwortung muss erhebliches hochschuldidaktisches Wissen herangezogen werden. *Wolff-Dietrich Webler* fixiert in seinem Beitrag **Formative und summative Evaluation schärfer unterscheiden! Anmerkungen über den Unterschied zwischen Rückmeldung und (summativer) Evaluation** die Probleme und bietet Antworten an.

Seite 48

W.W.

Werner Marx & Lutz Bornmann

Der Journal Impact Factor – Ein problematischer bibliometrischer Indikator



Werner Marx



Lutz Bornmann

Seit 1975 werden in den Journal Citation Reports von Thomson Reuters die Zeitschriften nach ihrem Journal Impact Factor (JIF) in Ranglisten eingestuft, welche jährlich aktualisiert werden. Zur Bestimmung der JIFs werden die Publikationen einer Zeitschrift jeweils aus einem Zeitfenster von zwei Jahren herangezogen und deren Zitierungen innerhalb des darauf folgenden Jahres ermittelt. Die Anzahl der Zitierungen wird dann durch die Anzahl der zitierbaren Publikationen (das sind Articles, Reviews und Short Communications) aus den beiden vorangegangenen Jahren dividiert. Inzwischen werden JIFs nicht nur zur Bewertung von Zeitschriften, sondern vielfach auch zur Bewertung einzelner Publikationen sowie zur Evaluierung von Forschern oder ganzer Forschungsinstitute herangezogen. In diesem Beitrag geben wir einen Überblick über häufig diskutierte kritische Einwände gegenüber dem JIF, alternative Metriken sowie die Grenzen seiner Anwendung im Hinblick auf den ursprünglichen Zweck (die Bewertung von Zeitschriften).

Grundsätzliche Bemerkungen

Seit 1975 werden in den Journal Citation Reports (JCR) wissenschaftliche Zeitschriften nach ihrem Journal Impact Factor (JIF) in Ranglisten eingestuft, welche jährlich aktualisiert werden (Thomson Reuters, 2012b). Wenn von Impact Factors die Rede ist, sind in der Regel diese JIFs von Thomson Reuters gemeint. Zunächst waren JIFs für die Einstufung der Zeitschriften als Ganzes im Zusammenhang mit deren kritischer Auswahl als Grundlage des Science Citation Index (SCI) vorgesehen. Inzwischen werden JIFs vielfach auch zur Bewertung einzelner Publikationen sowie zur Evaluierung von Forschern oder ganzer Forschungsinstitute herangezogen. Besondere Aufmerksamkeit kommt von Seiten der Zeitschriften-Verlage. Dementsprechend ist auch die Anzahl von Publikationen über diesen Indikator drastisch angestiegen. Die JIFs sind damit von einem zunächst wenig bekannten Hilfsmittel für Datenbankhersteller (und auch für Bibliothekare) zu einem weithin bekannten bibliometrischen Indikator zur Bewertung von Forschungsleistung geworden.

Die JCR werden in zwei Editionen angeboten: Die Zeitschriften der Science Edition umfassen derzeit rund 8000 Zeitschriften aus Naturwissenschaft, Medizin und Technik, die der Social Sciences Edition rund 2600 Zeit-

schriften aus den Geistes- und Sozialwissenschaften. Analog zu den unter dem Web of Science (WoS) bereitgestellten Zitierindexen von Thomson Reuters werden auch in den JCR nicht alle gegenwärtig erscheinenden Zeitschriften einbezogen, sondern lediglich eine strenge Auswahl, schätzungsweise nur 10% aller weltweit verfügbaren wissenschaftlichen Journale (die sogenannten Source Journals). Das ist jedoch für viele Belange ausreichend, da gezeigt werden kann, dass die Wissenschaft von nur relativ wenigen Zeitschriften bestimmt wird (Bradford's Law). Auswahlkriterien für die JCR sind die internationale Reputation der Editoren, die internationale Abdeckung des jeweiligen Fachgebietes, das regelmäßige Erscheinen, ein funktionierendes Qualitätssicherungsverfahren und weitere, zum Teil mehr subjektive Kriterien. Eine in den JCR nicht berücksichtigte Zeitschrift muss deshalb keine schlechte Zeitschrift sein.

Die aus der Definition des JIF resultierenden Fehlerquellen, die Probleme bei der Anwendung und Interpretation von JIFs sowie viele grundsätzliche Aspekte bibliometrischer Indikatoren sind Gegenstand zahlreicher Publikationen. Hier wird lediglich auf einige grundlegende und zusammenfassende Arbeiten verwiesen (Adler/Ewing/Taylor/Hall 2009; Amin/Mabe 2003; Bornmann/Marx/Gasparyan/Kitas 2012; Garfield 2006; Glänzel/Moed 2002; Kaltenborn/Kuhn 2003; Seglen 1997; Thomson Reuters 2012a; Vanclay 2012).

Definition und Bestimmungsverfahren

Zur Bestimmung der JIFs werden die Publikationen einer Zeitschrift jeweils aus einem Zeitfenster von zwei Jahren herangezogen und deren Zitierungen innerhalb des darauf folgenden Jahres ermittelt. Die Anzahl der Zitierungen wird dann durch die Anzahl der zitierbaren Publikationen (das sind Articles, Reviews und Short Communications) aus den beiden vorangegangenen Jahren dividiert. Zur Bestimmung beispielsweise des JIF-2010 wird zunächst die Anzahl der (zitierbaren) Publikationen aus den Jahren 2008 und 2009 festgestellt. Dann wird untersucht, wie oft der jeweilige Zeitschriftenname (Zeitschriftentitel) in den diversen zitierten Schreibweisen in Kombination mit den Publikationsjahren 2008 und 2009 in den Referenzen der Publikationen des Jahres 2010 erschienen ist. Schließlich wird der Quotient aus den gezählten Publikationen und deren Zitierungen gebildet. Etwas vereinfacht ausgedrückt kann man

sagen, dass der JIF einer Zeitschrift die mittlere Anzahl der Zitierungen innerhalb eines Jahres von Veröffentlichungen dieser Zeitschrift circa zwei Jahre nach deren Publikation angibt. Die JCR bieten inzwischen zusätzlich auch JIFs auf der Basis eines fünfjährigen (statt des klassischen zweijährigen) Publikationszeitraums an. So basiert z.B. der fünfjährige JIF-2010 auf den Publikationen aus den Jahren 2005 bis 2009 und deren Zitierungen im Jahre 2010. Da für die Berechnung des JIF nicht die kompletten Referenzen herangezogen werden, sondern nur die Häufigkeit des Vorkommens der Zeitschriftennamen in den Referenzen festgestellt wird, spricht man auch von Journal Title Matching. Dieses Verfahren hat den Vorteil, dass auch die verschriebenen Zitierungen berücksichtigt werden, da die Autorennamen und die numerischen Angaben der Referenzen unberücksichtigt bleiben. Alternativ könnte man auch unter den Referenzen aller gespeicherten Dokumente nachsehen, wie oft jede einzelne Publikation einer Zeitschrift aus einem zweijährigen Zeitintervall im darauf folgenden Jahr zitiert wurde. Der genaue Abgleich erfolgt normalerweise über den Namen des Erstautors, das Publikationsjahr, die Bandnummer sowie die Anfangsseite (Startseite) – der Zeitschriftenname wird mangels Vereinheitlichung nicht herangezogen. Dieses Citation Matching arbeitet jedoch wesentlich selektiver und liefert deutlich niedrigere Zitierungszahlen, da die verschriebenen Referenzen nicht mitgezählt werden. Anhand der Zitat-Varianten von hoch-zitierten Arbeiten kann man leicht zeigen, dass im Mittel immerhin rund 5% aller im SCI abgespeicherten Referenzen alleine bezüglich der darin enthaltenen Zahlenangaben verschrieben sind.

Fachunterschiede und Sonderfälle

Die verschiedenen Disziplinen und deren Fachkategorien zeichnen sich durch ganz unterschiedliche Publikations- und Zitiergewohnheiten bzw. -kulturen aus: Wenn man z.B. die Zitierungsraten (mittlere Anzahl der Zitierungen pro Publikation) in den verschiedenen Disziplinen vergleicht, findet man einen Unterschied von fast einem Faktor zehn: Mathematics und Computer Science weisen die niedrigste Zitierungsrate auf, Molecular Biology & Genetics die höchste und z.B. Chemistry sowie Physics liegen dazwischen (Thomson Reuters 2012a).

Ursache ist die unterschiedliche mittlere Anzahl von Referenzen im Anhang der Arbeiten aus den verschiedenen Disziplinen. Die Anzahl der potenziell zitierbaren Publikationen wie auch die unterschiedliche Politik der Zeitschriften spielen hier hinein. Die Fachunterschiede sind so groß, dass z.B. eine Zeitschrift an der Spitze des Rankings der einen Fachkategorie in einer anderen Kategorie mit dem gleichen JIF am Schluss des Rankings erscheinen kann. Außerdem muss berücksichtigt werden, dass die Zitiergewohnheiten innerhalb der einzelnen Fachgebiete auch zeitlichen Trends unterworfen sein können. Dementsprechend können also nur Zeitschriften miteinander verglichen werden, die zur gleichen Fachkategorie und zum gleichen JCR-Jahrgang gehören. Weiterhin unterscheiden sich Fachgebiete hin-

sichtlich der Geschwindigkeit des Publizierens. Publikationen dynamischer Gebiete, wie einige Biowissenschaften, erhalten einen verhältnismäßig hohen Anteil ihrer Zitierungen bereits in den ersten Jahren nach der Publikation und sind durch die JIF-Formel begünstigt.

Einige Disziplinen und insbesondere die Geistes- und Sozialwissenschaften leiden darunter, dass die Abdeckung der Publikationen durch die JCR-Zeitschriften völlig unzureichend sein kann (Archambault/Gagné 2004; Archambault/Vignola-Gagne/Cote/Lariviere/Gingrasb 2006). Denn die Wissenschaftler dieser Disziplinen publizieren oft nur einen kleineren Anteil ihrer Arbeiten in den erfassten Kernzeitschriften. Stattdessen erscheinen viele ihrer Publikationen in Büchern oder in Zeitschriften, die wegen ihrer engen fachlichen Ausrichtung, ihrer lediglich nationalen Bedeutung oder aus anderen Gründen nicht berücksichtigt werden. In geringerem Umfang gilt dies auch für die Technik-orientierten Disziplinen wie z.B. die Computer- und Materialwissenschaften. Die Begrenzung auf einen Satz von Kernzeitschriften impliziert nicht nur eine unvollständige Erfassung der zu bewertenden Publikationen (die Zitierungen der nicht erfassten Arbeiten fallen komplett weg), sondern schneidet auch einen Teil der die erfassten Arbeiten zitierenden Publikationen ab: Es fehlen alle Zitierungen durch zitierende Publikationen, deren Zeitschriften nicht berücksichtigt sind.

Problematik der Zeitfenster

Die klassische JIF-Formel beinhaltet ein relativ kurzes Zitierfenster (Zeitintervall, das bei der Zitationsanalyse berücksichtigt wird) und lässt nur eine mittlere Laufzeit von zwei Jahren zu (das Alter der Publikationen beträgt mindestens ein Jahr und maximal drei Jahre). Diese Festlegung resultiert vermutlich aus der Absicht, möglichst gegenwartsnahe JIFs anzubieten. Damit wird aber nur ein kleiner Ausschnitt der Zeitentwicklung der Zitierungen erfasst, der über die nach vielen Jahren (oder gar nach Jahrzehnten) gemessene Wirkung wenig aussagt. Außerdem liegt die Startzeit an verschiedenen Stellen des zweijährigen Publikationszeitfensters: Die am Anfang dieses Zeitfensters publizierten Arbeiten fallen in den Bereich der größten Resonanz, während bei den am Ende publizierten Arbeiten nur die Zitierungen während der Anlaufphase des ersten Jahres erfasst werden. Denn die Zitierungen der großen Masse der Arbeiten erreichen nach zwei bis drei Jahren ein Maximum und fallen dann mehr oder weniger schnell ab. Das gilt aber nur für die naturwissenschaftlichen Disziplinen – in manchen Disziplinen der Geisteswissenschaften wird das Maximum der Wirkung erst nach einem Jahrzehnt erreicht. Die Heranziehung des JIF auf der Basis fünfjähriger Publikationszeitfenster bietet diesbezüglich Vorteile und dämpft auch die jährlichen Schwankungen, doch das Publikationsalter und der Ausschnitt der Rezeptionskurve weisen eine größere Bandbreite auf. Außerdem ist dies mit einem Verlust an Aktualität verbunden: Der älteste berücksichtigte Jahrgang mit der längsten Laufzeit der Zitierungen und entsprechend größerem Gewicht liegt immerhin sechs Jahre zurück.

Weitere Verzerrungen und Fehlerquellen

Die in den JCR aufgenommenen Zeitschriften spiegeln die Dominanz und Qualität US-amerikanischer Forschung in vielen Fachgebieten wider, aber auch die Vorliebe der Wissenschaftler (nicht nur amerikanischer) für englischsprachige Zeitschriften. Man hat verschiedentlich vermutet, dass der insgesamt große Anteil amerikanischer Publikationen in der Forschung, speziell in dynamischen Forschungsgebieten wie einigen Biowissenschaften, sowie die Zitierung vorzugsweise von Kollegen des eigenen Landes die mittlere Anzahl Zitierungen amerikanischer Wissenschaftler wie auch den mittleren Journal Impact Factor amerikanischer Zeitschriften vergrößert. Doch die verschiedenen Möglichkeiten einer Begünstigung lassen sich nur schwer von einem Qualitätseffekt der Beiträge trennen. Die Bevorzugung inländischer Fachkollegen ist außerdem keine Form des Chauvinismus, sondern hängt mit der Sprache der Veröffentlichungen bzw. der Verfügbarkeit der Zeitschriften und persönlicher Bekanntschaft zusammen: Die Amerikaner finden unter den amerikanischen Arbeiten genügend zitierbare Publikationen.

Die Methode des Journal Title Matching ist leider eine nicht vernachlässigbare Quelle von Fehlern bei der Erstellung der JCR-Daten. Die Zeitschriftennamen bei Zitierungen sind nicht standardisiert und werden ohne Nachbearbeitung so in die Zitierdatenbanken übernommen, wie sie von den Autoren in Form der Referenzen angegeben wurden. Dadurch bedingt sind die Grenzen zwischen zwei Zeitschriften mit ähnlichen Namen fließend (z.B. *Chemical Physics* und *Journal of Chemical Physics*). Deshalb kann es vorkommen, dass JIFs mehr oder weniger stark verfälscht werden. Grundsätzlich kann man feststellen, dass die Wahrscheinlichkeit von Fehlern mit der Komplexität von Geschichte und Namen/Titeln von Zeitschriften ansteigt. Mit Zeitschriften, die diese Problematik aufweisen, scheinen die Datenbankhersteller bzw. deren Qualitätskontrolle vielfach überfordert zu sein. Fachübergreifende Zeitschriften wie *Nature*, *Science* oder *Proceedings of the National Academy of Sciences* (PNAS) sind in die Fachkategorie Multidisciplinary eingeordnet. Die darin erscheinenden Arbeiten stammen aus ganz unterschiedlichen Fachdisziplinen mit stark voneinander abweichenden Zitierkulturen (siehe oben). Da *Nature* und *Science* quantitativ von den Biowissenschaften dominiert werden und diese Disziplinen auch die höchsten Zitieraten aufweisen, wird auch der JIF dieser Zeitschriften entsprechend angehoben. Deren JIFs sind deshalb nicht mit denen von Zeitschriften anderer Disziplinen außerhalb der Biowissenschaften vergleichbar. Ein anderes Problem stellt die Mehrfach-Kategorisierung von Zeitschriften dar (die Zeitschrift *Scientometrics* wurde beispielsweise den Kategorien Computer Science und Interdisciplinary Applications zugeordnet) sowie die manchmal fragwürdige Zuordnung von Zeitschriften zu den JCR Fachkategorien (*Scientometrics* publiziert vor allem Beiträge aus dem Bereich der Messung von Forschungsleistungen, aber kaum aus der Informatik).

Citing/Cited Half Life und Immediacy Index

Die JCR liefern neben den JIFs noch drei weitere interessante Kenngrößen für das Zitierverhalten im Hinblick auf Zeitschriften, die von Thomson Reuters wie folgt definiert sind:

- Citing Half-Life = Anzahl zurückliegender Jahre seit dem gewählten Jahr, aus denen 50% aller in den Arbeiten einer Zeitschrift des betreffenden Jahres zitierten Publikationen stammen. Hier geht es also um die Referenzen der Artikel einer Zeitschrift.
- Cited Half-Life = Anzahl zurückliegender Jahre seit dem gewählten Jahr, aus denen 50% aller im betreffenden Jahr zitierten Arbeiten einer Zeitschrift stammen. Hier geht es also um die Arbeiten einer Zeitschrift als Referenzen im Anhang der Publikationen anderer Zeitschriften.

Die Citing Half-Life macht demnach eine Aussage über das Zitierverhalten der Autoren einer Zeitschrift und spiegelt deren Auswahlverhalten beim Zitieren wider. Die Altersverteilung der Referenzen informiert über die Gegenwartsnähe (Aktualität) der zitierten Literatur, auf der die Artikel einer Zeitschrift aufbauen. Die Cited Half-Life informiert dagegen über die Zitierungen der Artikel einer Zeitschrift durch andere Arbeiten und gibt wieder, wie lange man sich an die Artikel dieser Zeitschrift erinnert (d.h. über die Nachhaltigkeit ihrer Wirkung, was durchaus auch als ein Qualitätskriterium aufgefasst werden kann). Cited und Citing Half-Life machen beide eine Aussage über die Abnahme des Interesses an Publikationen (das Altern), denn dieselben Publikationen bzw. Autoren zitieren andere und werden von anderen zitiert. Die Citing Half-Life ist auf die Arbeiten einer bestimmten Zeitschrift bezogen aktiv, die Cited Half-Life dagegen passiv.

- Immediacy Index (II) = Anzahl Zitierungen der Arbeiten einer Zeitschrift im Jahr der Publikation dividiert durch die Anzahl der (zitierbaren) Publikationen dieser Zeitschrift im gleichen Jahr.

Der II-2010 einer bestimmten Zeitschrift errechnet sich demnach aus der Anzahl Zitierungen im Jahr 2010 der Publikationen dieser Zeitschrift aus dem gleichen Jahr, dividiert durch die Anzahl der Publikationen wiederum im Jahr 2010. Der II kann als eine Sonderform des JIF angesehen werden, bei dem weniger die Resonanz insgesamt, sondern vielmehr die Schnelligkeit der Reaktion der Fachkollegen gemessen wird. Einschränkend muss man jedoch feststellen, dass die jeweils am Ende des Jahres erschienenen Publikationen noch kaum eine Chance hatten, Zitierungen zu akkumulieren und deshalb immer unterbewertet werden.

Grenzen der Anwendung

Die Tatsache, dass eine bestimmte Arbeit in einer Zeitschrift mit hohem oder niedrigem JIF erschienen ist, sagt wenig über den Impact dieser Arbeit aus. Denn selbst in Zeitschriften mit hohem JIF hat die Masse der Arbeiten keinen ungewöhnlich hohen Impact. Dementsprechend wird eine typische (d.h. eine um den Median der Zitie-

rungen liegende) Arbeit einer Zeitschrift mit hohem Impact nicht wesentlich häufiger zitiert, als eine typische Arbeit einer Zeitschrift mit niedrigem Impact. Die unsymmetrische bzw. schiefe Verteilung der Zitierungen auf die Publikationen (Pareto-Verteilung) bedingt, dass nur ein kleiner Anteil hoch-zitierter Arbeiten einen großen Anteil der Zitierungen erhält (Seglen 1992). Die meist-zitierte Hälfte der Arbeiten wird um ein Vielfaches häufiger zitiert als die verbleibende Hälfte. Der arithmetische Mittelwert (und damit der JIF einer Zeitschrift) wird also nur von relativ wenigen Arbeiten bestimmt und erlaubt deshalb keinen Rückschluss auf den Impact einer speziellen in dieser Zeitschrift publizierten Arbeit. Die Schiefe einer Verteilung kann mathematisch beschrieben und quantifiziert werden (z.B. über den Gini-Index). Man kann zeigen, dass die Schiefe der Zitierteilung einer Zeitschrift mit der Höhe ihres JIFs ansteigt. Dementsprechend sagt ein hoher JIF besonders wenig über den Impact der Masse der darin erschienenen Arbeiten aus. Ein extremes Beispiel ist die Zeitschrift *Acta Crystallographica A*, deren JIF innerhalb eines Jahres wegen einer einzelnen hoch-zitierten Arbeit von 2,051 (JCR-2008) auf 49,926 (JCR-2009) anstieg (Dimitrov/Kaveri/Bayry 2010). Wir haben es hier also mit einer Art von gerichteter Korrelation zu tun: Der Impact von einzelnen Arbeiten verhilft einer Zeitschrift zu mehr oder weniger Ansehen, die Publikation einer Arbeit in einer Zeitschrift mit hohem JIF bewirkt jedoch nicht, dass diese auch automatisch hoch zitiert wird. Deshalb ist es nicht verwunderlich, dass die zwischen den Zitierungen der Arbeiten eines einzelnen Wissenschaftlers und den JIFs der Zeitschriften dieser Arbeiten untersuchte Korrelation schwach ausfällt (Seglen 1997). JIFs sind also zur Feststellung des Impacts der Arbeiten eines Wissenschaftlers eher ungeeignet. Um den Impact einzelner Arbeiten festzustellen, sollte man die Zitierungen jeder einzelnen Publikation ermitteln, auch wenn die Recherche und Interpretation dieser Daten mit höherem Aufwand und mit (anderen) Fallstricken behaftet ist. Der Umweg der Impact-Bestimmung über die JIFs macht keinen Sinn, auch wenn in einzelnen Disziplinen (wie z.B. der Medizin) oder in verschiedenen Ländern (z.B. in skandinavischen Ländern, der Volksrepublik China und Indien) die Forschungsbewertung und Mittelvergabe gerne danach bemessen wird.

Diskussion

Die Erstellung von Zitationsdaten zu einzelnen Publikationen kann verhältnismäßig aufwendig sein (wenn man z.B. Zitierungen für ein fixes Zitierfenster recherchiert). Außerdem können neuere Arbeiten bzw. jüngere Wissenschaftler kaum über Zitierungen bewertet werden. Nicht zuletzt aus diesen Gründen wird vielfach auf die JIFs der Zeitschriften zurückgegriffen, in denen die Arbeiten erschienen sind. Die leichte Verfügbarkeit der JIF-Daten verleitet Forschungsplaner, die vermeintlich objektive und leicht nachvollziehbare Daten zur Evaluierung der Forschung fordern, zum leichtfertigen Umgang mit diesen Daten. Fehl- und Überinterpretationen von Ergebnissen zur Forschungsleistung sind die Folge.

Nachfolgend möchten wir die wichtigsten Aspekte zusammenfassen, die beim Einsatz von JIFs berücksichtigt werden sollten:

- Die Korrelation zwischen dem Impact einer einzelnen Arbeit und dem JIF der publizierenden Zeitschrift ist in der Regel gering
- Zeitschriften unterschiedlicher Fachgebiete sollten nicht über ihre JIFs miteinander verglichen werden,
- Dynamische Fachgebiete, wie einige Biowissenschaften, akkumulieren ihre Zitierungen schneller und werden durch die JIF-Formel bevorzugt.
- Einzelne Disziplinen, wie die Mathematik oder die Geistes- und Sozialwissenschaften, publizieren große Teile Ihrer Arbeiten in Büchern.
- Fachübergreifende Zeitschriften, wie *Nature* oder *Science*, werden durch die großen Anteile von nicht-klassischen Artikeln aufgewertet.
- Änderungen von Zeitschriftennamen (Zeitschriftentiteln) beeinflussen die JIFs noch mehrere Jahre nach der Namensänderung.
- Die aktuellsten JIFs bewerten die Publikationen der beiden vorangegangenen Jahre und nicht des jüngsten Jahrgangs einer Zeitschrift.

JIFs sind mit Einschränkung durchaus geeignet, das aktuelle Ansehen von Zeitschriften als Ganzes und den Einfluss innerhalb ihrer Fachdisziplinen einzuschätzen. Eine absolute und genaue Maßzahl für deren Qualität oder gar für die Qualität einer einzelnen darin erschienenen Arbeit bzw. der Arbeiten eines dort publizierenden Wissenschaftlers sind sie jedoch nicht. Der Hersteller Thomson Reuters rät auch im Falle der JIFs (wie allgemein beim Einsatz von Zitierungen zur Forschungsevaluierung), diese vorsichtig und nur in Kombination mit anderen Indikatoren heranzuziehen. JIFs sind, wie andere bibliometrische Indikatoren auch, nicht als Ersatz für das Expertenvotum, sondern als ein zusätzlicher unabhängiger Indikator anzusehen. Die Frage ist weniger, ob, sondern vielmehr wie man den JIF gebraucht.

Literaturverzeichnis

- Adler, R./Ewing, J./Taylor, P./Hall, P.G. (2009): A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS): [Article]. In: *Statistical Science*, 24(1), pp. 1-28. doi: 10.1214/09-sts285.
- Amin, M./Mabe, M.A. (2003): Impact factors: use and abuse. In: *Medicina-Buenos Aires*, 63(4), pp. 347-354.
- Archambault, E./Gagné, É.V. (2004): *Science-Metrix: the use of bibliometrics in the Social Sciences and Humanities*. Montreal, Canada: Social Sciences and Humanities Research Council of Canada (SSHRC).
- Archambault, E./Vignola-Gagne, E./Cote, G./Lariviere, V./Gingras, Y. (2006): Benchmarking scientific output in the social sciences and humanities: The limits of existing databases. In: *Scientometrics*, 68(3), pp. 329-342.
- Bornmann, L./Marx, W./Gasparyan, A.Y./Kitas, G.D. (2012): Diversity, value and limitations of the Journal Impact Factor and alternative metrics. In: *Rheumatology International (Clinical and Experimental Investigations)*, 32(7), pp. 1861-1867. doi: 10.1007/s00296-011-2276-1
- Dimitrov, J. D./Kaveri, S. V./Bayry, J. (2010): Metrics: journal's impact factor skewed by a single paper. In: *Nature*, 466(7303), pp. 179. doi: 10.1038/466179b.
- Garfield, E. (2006): The history and meaning of the Journal Impact Factor. In: *Journal of the American Medical Association*, 295(1), pp. 90-93.
- Glänzel, W./Moed, H. F. (2002): Journal impact measures in bibliometric research. In: *Scientometrics*, 53(2), pp. 171-193.

- Kaltenborn, K.F./Kuhn, K. (2003): Der Impact-Faktor als Parameter zur Evaluation von Forscherinnen/Forschern und Forschung. In: Medizinische Klinik, 98(3), pp. 153-169.
- Reuters, T. (2012a): Essential Science Indicators. Retrieved January 12, 2012, from http://thomsonreuters.com/products_services/science/science_products/a-z/essential_science_indicators/
- Reuters, T. (2012b): Journal Citation Reports. Retrieved January 12, 2012, from <http://scientific.thomson.com/products/jcr/>
- Seglen, P.O. (1992): The skewness of science. In: Journal of the American Society for Information Science, 43(9), pp. 628-638.
- Seglen, P.O. (1997): Why the impact factor of journals should not be used for evaluating research. British Medical Journal, 314(7079), pp. 498-502.
- Vanclay, J. K. (2012): Impact Factor: outdated artefact or stepping-stone to journal certification? In: Scientometrics, 92(2), pp. 211-238.

- **Werner Marx**, Max-Planck-Institut für Festkörperforschung, Informationsvermittlungsstelle (IVS-CPT), E-Mail: w.marx@fkf.mpg.de
- **Lutz Bornmann**, Generalverwaltung der Max-Planck-Gesellschaft, Stabsreferat für Wissenschafts- und Innovationsforschung, Forschungsanalyse, E-Mail: bornmann@gv.mpg.de

**Wolff-Dietrich Webler (Hg.):
Universitäten am Scheideweg ?! - Chancen und Gefahren des gegenwärtigen historischen Wandels in Verfassung, Selbstverständnis und Aufgabenwahrnehmung
Ergebnisse des Hochschulforums Sylt 2008**

Ist der Weg von der Idee der Gemeinschaft der Lehrenden und Lernenden zu Universitäten in differenzierten Leistungsklassen als Produktionsunternehmen für wirtschaftlich verwertbare Erkenntnisse und hoch qualifizierte Arbeitskräfte unumkehrbar? Gibt es einen dritten Weg?

Die Entwicklung hat sich schon Jahrzehnte abgezeichnet – jetzt ist der Wandel in vollem Gange (und vermutlich unumkehrbar). Die Universitätsleitungen in Deutschland sehen sich – von ihnen gewollt oder nicht – einer Entwicklung gegenüber, die "ihre" Universität täglich verändert und die – provokant zugespitzt – in die Formel gefasst werden kann:

Von der Idee der Gemeinschaft der Lehrenden und Lernenden in grundsätzlich gleichen (gleichrangigen) Universitäten zu einem Produktionsunternehmen in differenzierten Leistungsklassen, das Wirtschaftlichkeitsregeln durchgängig folgt und das vordringlich wirtschaftlich verwertbare Erkenntnisse und Arbeitskräfte erzeugt.

Diese Situation, die die deutsche Universität so nachhaltig verändern wird wie kaum etwas anderes vorher, stand im Zentrum des Hochschulforums Sylt 2008. Dort wurde gefragt:

Gibt es einen dritten Weg?

Die zentrale These lautet: Wenn nicht korrigierend eingegriffen wird, dann wird die Universität als kollegiale Veranstaltung verlassen – mit weitreichenden Folgen für Zusammenhalt, Produktivität, Verantwortungsstrukturen, für Art, Niveau und Profil von Forschung, Lehre und Studium bzw. Art, Niveau und Profil der Absolvent/innen. Bisherige kollegial integrative Meinungsbildungs-, Entscheidungs-, personelle Ergänzungs-(Berufungs-)verfahren werden von betriebsförmigen Strukturen abgelöst. Dieses Neue enthält Chancen und Gefahren – in welchem Umfang und mit welchem Ergebnis ist offen. Das Ergebnis aber ist für die deutsche Gesellschaft und weit darüber hinaus von allergrößter Bedeutung. Hier setzt das in diesem Band vorgelegte Konzept des Hochschulforums 2008 an.

Hochschulforscher, Universitätsrektoren/-präsidenten und Mitglieder aus Wissenschaftsministerien haben sich für acht Tage in Klausur begeben, mit dem Ziel die weiteren Konsequenzen der Maßnahmen zu vergegenwärtigen und sich zu vergewissern, ob und wie diese Folgen gewollt werden.

Das Ergebnis – bestehend aus Analysen und Handlungsempfehlungen – wird hiermit vorgelegt.



ISBN 3-937026-64-9, Bielefeld 2009,
296 Seiten, 39.80 Euro

Mit Beiträgen von:

Philip G. Altbach, Tino Bargel,
Hans-Dieter Daniel, Christiane Gaehtgens,
Ludwig Huber, Wilhelm Krull,
Stephan Laske, David Lederbauer,
Bernadette Loacker,
Claudia Meister-Scheytt,
Klaus Palandt, Ulrich Peter Ritter,
Thomas Rothenfluh, Christoph Scherrer,
Jürgen Schlegel, Boris Schmidt,
Dieter Timmermann, Carsten von Wissel,
Wolff-Dietrich Webler, Gülsan Yalcin,
Frank Ziegele.

Bestellung - E-Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

Heinrich Seidlmeier



Heinrich Seidlmeier

Modellierung von Qualitätsaspekten in Hochschulprozessen

Insbesondere auf den zahlreichen Kongressen zum Thema „Qualitätsmanagement an Hochschulen“ wird von den Referenten häufig auf sogenannte „Prozesslandkarten“ verwiesen (vgl. zum Qualitätsmanagement an Hochschulen z.B. Winde 2010). Diese enthalten, meist in ansprechender, übersichtlicher Form, die identifizierten Hochschulprozesse; in der Regel auf der ersten Gliederungsebene gruppiert in Kern-, Unterstützungs- und Führungsprozesse. Gleichzeitig wird die Bedeutung der Prozesssicht für das Qualitätsmanagement betont. Dagegen ist nichts einzuwenden – ganz im Gegenteil (vgl. zum prozessorientierten Qualitätsmanagement an Hochschulen z.B. von Lojewski 2009).

Kommt die Sprache jedoch auf einzelne, konkrete lehr- oder forschungsnahe Prozesse auf den unteren Gliederungsebenen, wie Neuberufungen, Einrichtung von neuen Studiengängen, Durchführung von Akkreditierungen, Evaluierungen oder Prüfungen ist gewöhnlich schnell ersichtlich, dass nur rechtliche Vorgaben „nachmodelliert“ wurden. Dagegen ist grundsätzlich auch noch nichts zu sagen. Nur, Rechtsverordnungen wie Hochschulpersonalgesetze der Länder oder Grundordnungen der Hochschulen alleine können die Qualität in Lehre und Forschung nicht sichern. Selbst wenn in einem Neuberufungsverfahren – dieses Beispiel durchzieht diesen Artikel – alle rechtlichen Aspekte genauestens eingehalten werden, ist nicht sichergestellt, dass am Ende die oder derjenige Beste berufen wird, der tatsächlich Lehre auf höchstem Qualitätsniveau im Hörsaal bieten wird.

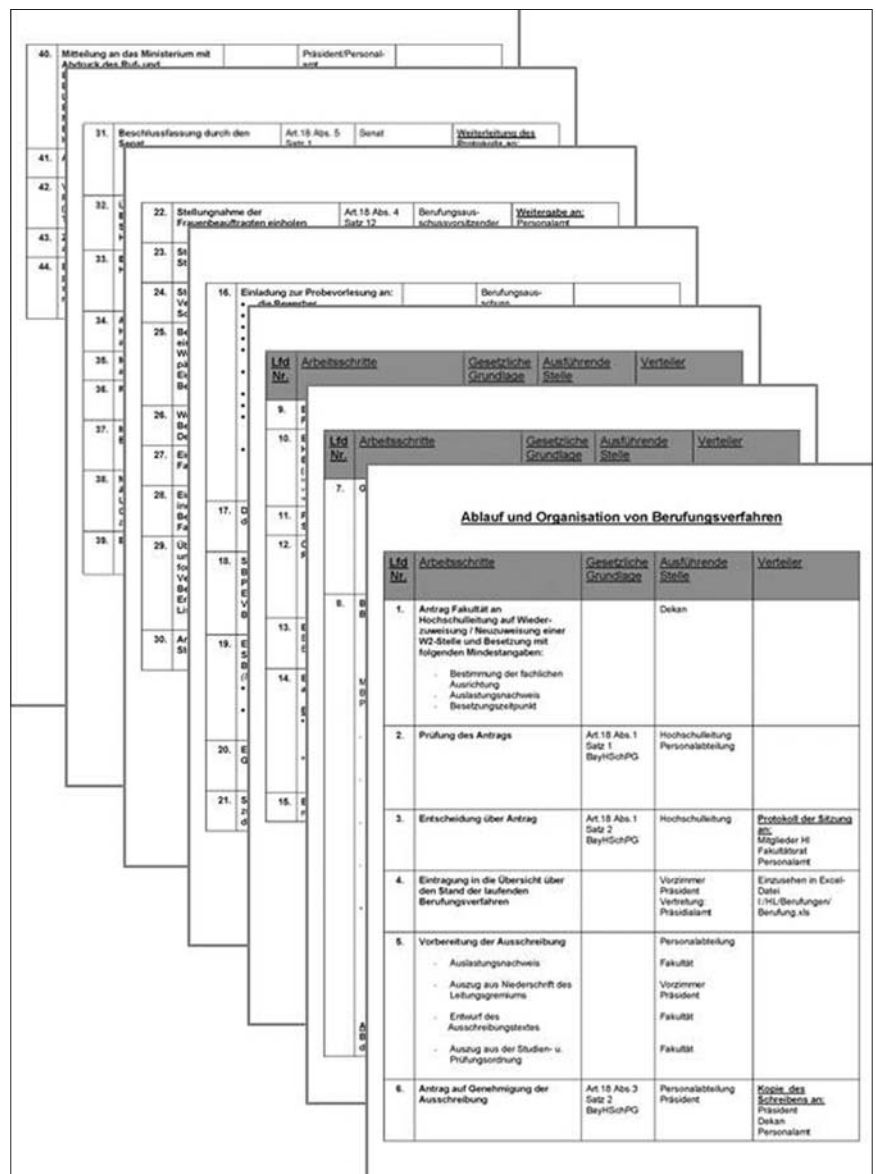
Dieser Artikel soll in kurzer Form zeigen, wie Modelle für Hochschulprozesse mit qualitätskritischen Aspekten angereichert werden können.

1. Hochschulprozesse in Listenform

Um die Anschaulichkeit der Problematik zu erhöhen, wird als Grundlage für die weiteren Ausführungen der

Prozess „Neuberufung von Professor/innen“ gewählt. Zweifelsohne beeinflusst das Ergebnis dieses Prozesses, die oder der Neuberufene, ganz maßgeblich die Qualität¹ im ausgeschriebenen Lehr- und Forschungsgebiet.

Abbildung 1: Textuelle Beschreibung Neuberufungsverfahren



¹ Der Begriff der Lehr- bzw. Forschungsqualität soll an dieser Stelle nicht thematisiert werden.

Versteht man unter einem Kernprozess (oder Geschäftsprozess) einen wertschöpfenden oder auch hochschulstrategisch wichtigen Prozess, handelt sich bei der Auswahl und Einstellung neuer Kolleg/innen sicherlich um einen Prozess, der den Kern des Hochschulgeschäfts im Innersten betrifft.

Ist ein neues Berufungsverfahren einzuleiten, besteht die erste Aufgabe des Leiters der Berufungskommission mit darin, sich einen Überblick über den durchaus umfangreichen Ablauf zu verschaffen. Da Berufungen nicht zu den Routinen eines Hochschullehrers zählen, ist man hierzu meist für jede Hilfe dankbar. Gebräuchlich und durchaus nützlich sind textuelle Ablaufbeschreibungen oder „Checklisten“, die sogar als Grundlage eines „Projektplans“ verwendet werden können. Ein typisches Beispiel zeigt Abbildung 1.

Diese Beschreibung weist die nummerierten Arbeitsschritte, die zugehörigen gesetzlichen Grundlagen, die ausführenden Stellen sowie „Verteiler“ (Empfänger von

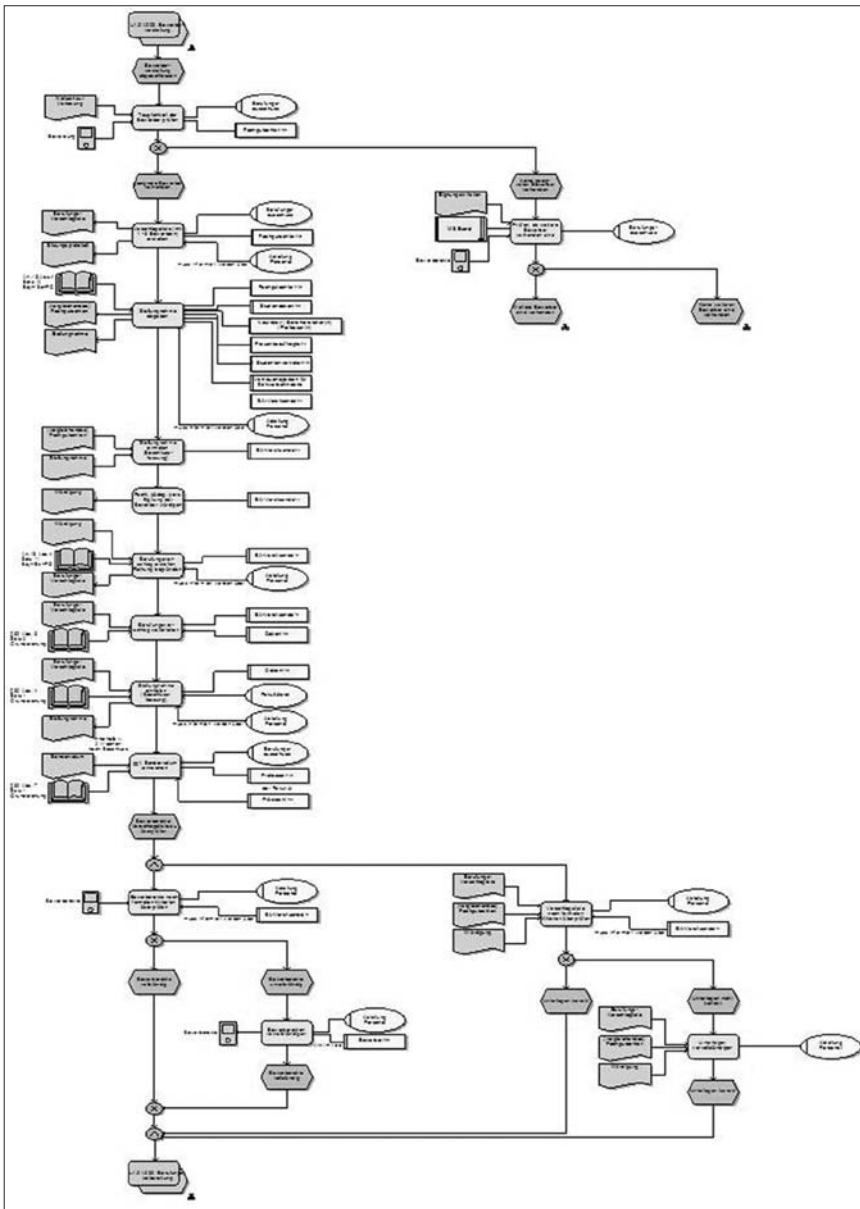
Unterlagen und Kopien sowie Ablageorte) auf. Auf die Komplexität des Prozesses verweist die mit 44 sehr hohe Anzahl an Prozessschritten hin.

Solche Listen haben eine Reihe von Nachteilen. Zum einen stellen sie keine vollständigen Prozessbeschreibungen dar (es fehlt im Wesentlichen die Ablauflogik). Wie bei Papierdokumenten allgemein üblich, ist die Änderung, Speicherung und Verteilung bzw. der Zugriff aufwändiger als bei elektronischen Unterlagen. Zum anderen können keine zusätzlichen Dokumente (wie Gesetzestexte) einfach angehängt oder verlinkt werden. Softwaregestützt erzeugte Prozessmodelle vermeiden diese Schwachstellen.

2. Hochschulprozesse in Modellform

Die Umsetzung von vorliegenden Listen in methodisch einwandfreie Prozessmodelle ist vergleichsweise einfach durchzuführen. Meist ergeben sich bei der Modellierung Ungenauigkeiten und Unvollständigkeiten, die in „Interviews“ mit der Personal- und Rechtsabteilung bzw. weiteren Prozessbeteiligten geklärt werden können.

Abbildung 2: Modellausschnitt zum Neuberufungsprozess (in EPK-Darstellung)



Im Ergebnis – meist ist man bemüht sehr genau und detailliert zu modellieren – entstehen oft für Experten sehr aussagekräftige und interessante Modelle, die in ihrer „Tapetenform“ Laien (wie Vorsitzende von Berufungsausschüssen oder Fremdgutachter) überfordern. Der Ruf nach der „guten, alten Liste“ wird spätestens bei der Präsentation der Modelle laut. Abbildung 2 zeigt einen Ausschnitt in Form einer „Ereignisgesteuerten Prozesskette“ (kurz: EPK; diese Modellierungsmethodik ist seit langen Jahren in Wissenschaft und Praxis weit verbreitet) (vgl. zur ARIS-Methodik und zum entsprechenden ARIS-Modellierungstool z.B. Seidlmeier 2010).

Vollständig modelliert umfasst der Neuberufungsprozess beispielsweise 77 Schritte (statt der 44 oben genannten) und 30 Prozessbeteiligte. Moderne Modellierungstools bieten Funktionen zur Handhabung von komplexen Modellen, indem sie zum Beispiel passende Überblickmodelle anbieten oder es erlauben, sehr lange Prozessketten in verknüpfte Teilprozesse zu segmentieren.

Die umfangreiche Modellierung von Prozessen ist selbstredend kein Selbstzweck (auch wenn grundsätzlich immer die Gefahr droht, zu ausführlich zu modellieren und die Möglichkeiten der vorhandenen Softwarewerkzeuge mehr als auszunutzen). Zudem muss dem Zusatzaufwand der Modellierung auch ein Zusatznutzen gegenüberstehen – zum Beispiel:

Der qualitätskritische Kern des Neuberufungsprozesses in der Ablaufliste der

Abbildung 3: Modellausschnitt „Bewerberbegutachtung“

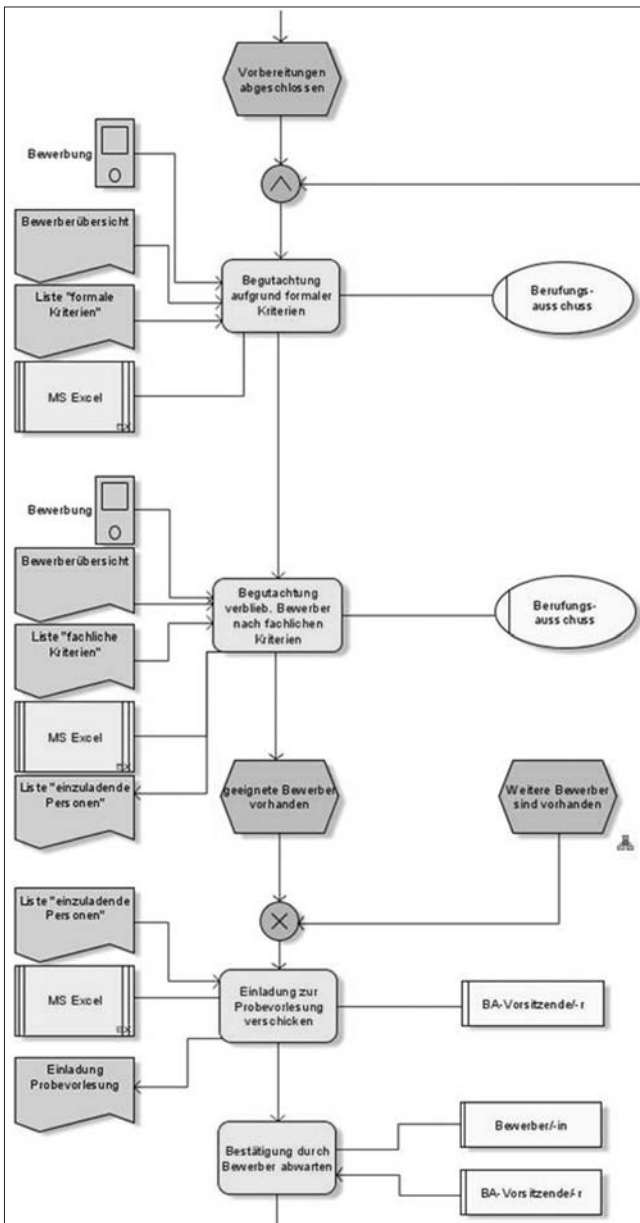


Abbildung 1 umfasst lediglich vier knapp beschriebene Schritte:

- Begutachtung der Bewerbungen aufgrund formaler Kriterien (Schritt 14 von insgesamt 44),
- Begutachtung der Bewerbungen nach fachlichen Kriterien (Schritt 15),
- Einladung zur Probevorlesung (Schritt 16),
- Durchführung und Auswertung der Probevorlesungen (Schritt 17).

Als ausführende Stelle ist für alle Schritte der Berufungsausschuss genannt sowie keine gesetzlichen Grundlagen und keine Verteiler. Unmittelbar einsichtig ist, dass mit dieser knappen Anweisung keinerlei Prozessqualität im Sinne von „Auswahl des/der besten Kandidat/in“ gewährleistet werden kann.

Die ansonsten sehr praktische Verfahrensanweisung in Listenform scheitert in dieser kritischen Phase, weil u.a. keine unterstützenden gesetzlichen Hinweise vorliegen (man kann davon ausgehen, dass solche Listen in der Regel in Personalabteilungen auf Basis rechtlicher Vorgaben entstehen).

Auch wenn die Begutachtung von Kandidaten nicht vollständig formalisierbar erscheint, können jedoch angebotene Hilfsmittel – wie in der folgenden Abbildung 3 auf der linken Seite modelliert – Hilfe für eine einheitliche, nachvollziehbare Bewertung durch mehrere Beteiligte bieten und damit das Prozessergebnis qualitativ verbessern. Wenn solche Hilfsmittel dem Berufungsausschuss nicht bekannt sind (auch wenn sie „irgendwo“ sogar vorliegen), weil die „Liste“ nicht darauf verweist (oder, weil sie nicht gepflegt wird, nur auf veraltete Dokumente), erzeugen Prozessmodelle mit diesen Hinweisen einen deutlichen Mehrwert. Selbstverständlich könnten auch Listen diese Verweise aufnehmen. Derartige Lücken und ihre Lösungen werden meist jedoch erst bei der genauen Ablaufmodellierung gesehen.

Im Gesamtprozess sind noch eine ganze Reihe weiterer Schritte qualitätskritisch. Schon die Gestaltung des Ausschreibungstextes und – ganz naheliegend – die Bildung des Berufungsausschusses und die Benennung der Fremdgutachter können (mit-)entscheidend sein. Schon die informative Anreicherung von Prozessmodellen ist also wertvoll. Man kann aber noch einen Schritt weitergehen.

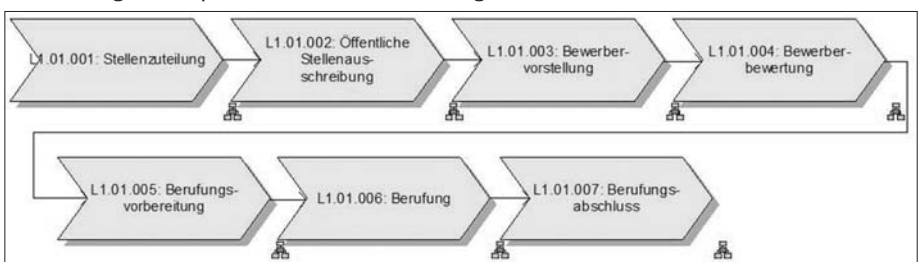
3. Qualitative Anreicherung von Prozessmodellen

Die Modellierung von Qualitätsaspekten, aber auch von Qualitätsrisiken, muss jedoch über die erwähnte Abbildung von nützlichen Listen, Excel-Tabellen u.Ä. sowie über die Erhöhung der Anzahl dargestellter Prozessschritte – und damit der Genauigkeit der Ablaufdarstellung – hinausgehen (im Beispiel von 44 auf 77).

An der Hochschule Rosenheim wurde zu diesem Zweck ein Kreis an neuberufungserfahrenen Kollegen aus allen Fakultäten zusammengerufen, um systematisch nach qualitätssteigernden Aspekten zu suchen. Als unterstützende „Systematik“ wurde das Prozessüberblicksmodell in Abbildung 4 zugrunde gelegt (im Workshop konnte somit das „Monster“ Neuberufungsprozess kontrolliert untersucht werden). Als weitere Grundlage dienten Erkenntnisse aus den modellgestützten Prozessanalysen „am Schreibtisch“.

Im Ergebnis entstanden in einer Sitzung 38 qualitätssichernde oder -fördernde Ideen, Ansätze und konkrete Maßnahmen, welche nur noch in einem anschließenden

Abbildung 4: Teilprozesse im Neuberufungsverfahren

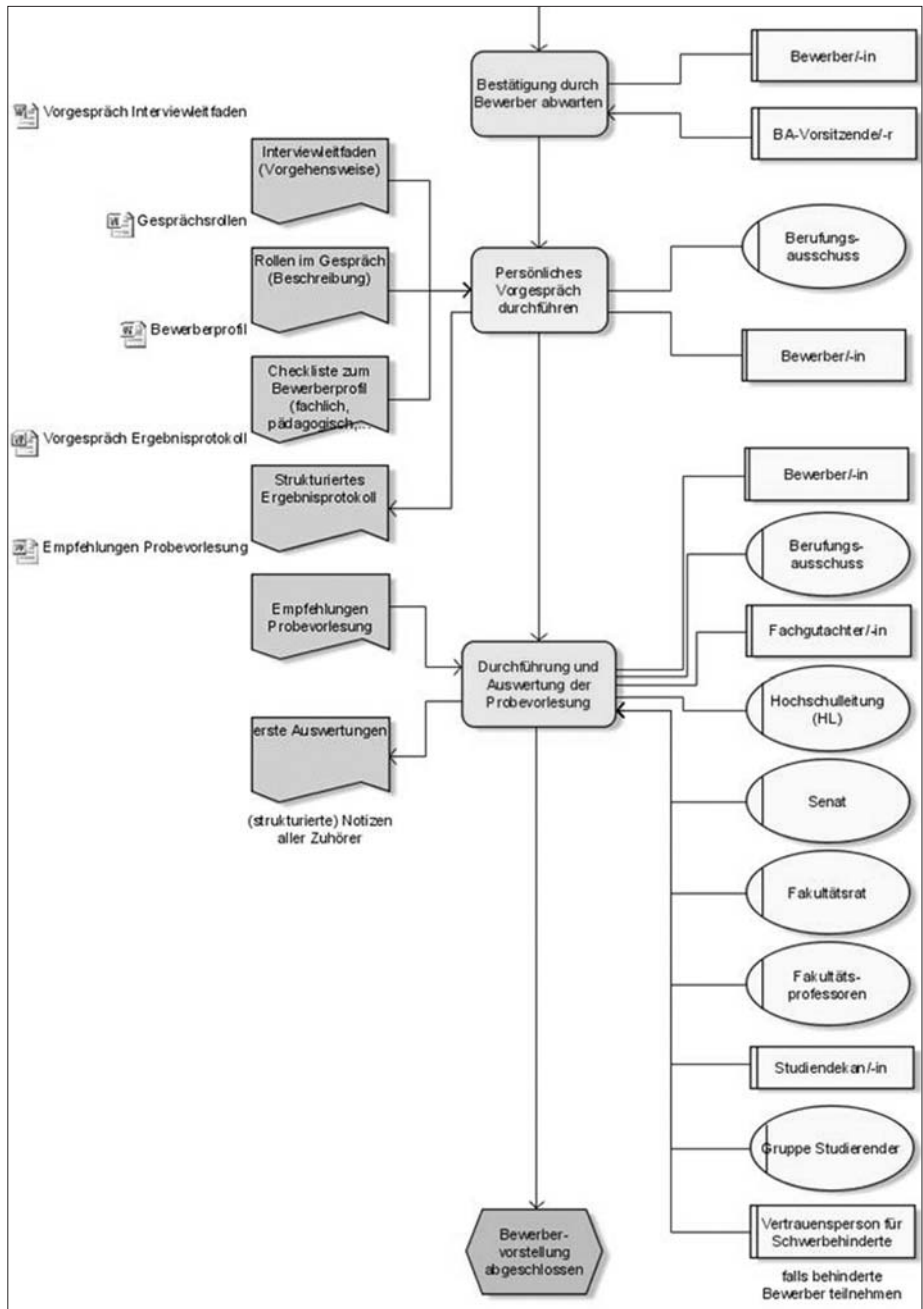


Rundlauf ohne erneutes Treffen der Kollegen geprüft und ergänzt werden mussten. Diese gleichermaßen sehr effiziente und effektive Vorgehensweise ist ohne Zweifel auch der konsequent prozessorientierten und dadurch disziplinierenden Denkweise geschuldet. Diese Qualitätsaspekte sind in die Modelle einzubeziehen.

Die qualitative Anreicherung des Prozessmodells mit Erfahrungswissen soll an einer beispielhaften Stelle erläutert werden. In der folgenden Abbildung 5 ist ersichtlich, dass der Schritt „Durchführung und Auswertung der Probevorlesung“ sehr personalintensiv ist.

Insgesamt sind zehn organisatorische Einheiten, Stellen und Personen beteiligt. Jede Probevorlesungsrunde ist damit extrem ressourcenintensiv. Die Wahl unpassender Kandidaten/innen im Vorfeld ist deswegen unbedingt zu vermeiden. Eine Lösung, den Aufwand und auch das Qualitätsrisiko (einer Fehlbesetzung) zu vermeiden, kann in der Einführung persönlicher Vorgespräche zwischen dem Berufungsausschuss und den Bewerbern/innen liegen. Werden im Prozessmodell einschlägige qualitätssichernde oder -fördernde Aspekte mit ausgewiesen, ist für die Mitglieder des Berufungsausschusses in der nicht alltäglichen Gesprächssituation ein erheblicher Mehrwert verbunden. Im Beispiel sind für den Schritt „Persönliches Vorgespräch führen“ Erfahrungen aus früheren Berufungsprozessen dokumentiert (Interviewleitfaden, Rollen usw.) und im Modell als Worddateien zugriffsbereit hinterlegt (auch als einfaches Dokumenten- bzw. Wissensmanagement zu verstehen, welches beispielsweise auch zu einem „Berufungs-Wiki“ ausgebaut werden könnte²).

Abbildung 5: Vorgespräche im Neuberufungsverfahren



4. Verwendung und Akzeptanz von Prozessmodellen

Abschließend ist noch auf zwei Punkte hinzuweisen. Zum einen arbeiten nur Modellierungsexperten mit den Softwaretools, die Prozessmodelle wie die in den Abbildungen 2 bis 5 erzeugen. Prozessbeteiligte (vom Hochschulpräsidenten bis zum Mitarbeiter in der Personalabteilung oder Professoren in Fakultäten) haben diese Werkzeuge im Normalfall nicht zur Verfügung und müssen bzw. wollen damit überhaupt nicht arbeiten. Beteiligte als Verwender der Modelle brauchen andere Zugriffsmöglichkeiten.

Zum anderen sind die abgebildeten Prozessmodelle für nicht geübte Laien (wie Vorsitzende von Berufungsausschüssen) ungewohnt und schwer lesbar. Sie erzeugen damit – gegenüber gebräuchlichen „Listen“ – einen gewissen Zusatzaufwand. Diese Modelle müssen zu ihrer Akzeptanz einen (weiter oben schon angesprochenen) Zusatznutzen erzeugen. Zwei Arten sind insbesondere zu nennen:

- Anreicherung mit erfahrungsbasierten Qualitätshinweisen,
- Verlinkungen und damit einfacher Zugriff auf immer aktuelle elektronische Dokumente.

² Unter einem „Wiki“ ist ein meist einfaches, webbrowsers-basiertes Informationssystem zu verstehen, welches weitgehend von seinen selbst Benutzern mit Informationen versorgt wird.

Zuletzt müssen zur Akzeptanz und damit zur Verwendung der Modelle zwei Voraussetzungen erfüllt sein. Die in der Grundform komplexen Prozessmodelle müssen durch eine rollenbasierte Personalisierung vereinfacht werden (die durch den Studiendekan auszuführenden Schritte sind für den Fremdgutachter uninteressant und zudem störend). Modelle sind digital so zu veröffentlichen, dass sie über einen Standard-Webbrowser einfach zu erreichen sind (im Sinne eines digitalen Prozesshandbuches). Das Abrufen der Modelle muss zusätzlich weitestgehend intuitiv möglich sein. Zeit- und akzeptanzraubende Schulungen und Softwareinstallationen sind so zu vermeiden.

Literaturverzeichnis

Lojewski v., U. (2009): Prozessorientiertes Qualitätsmanagement (an der Fachhochschule Münster). In: Richthofen v., A./Lent, H. (Hg.): Qualitätsentwicklung in Studium und Lehre. Bielefeld.
 Seidlmeier, H. (2010): Prozessmodellierung mit ARIS. 3. Auflage. Wiesbaden.
 Winde, M. (2010): Von der Qualitätsmessung zum Qualitätsmanagement - Praxisbeispiele an Hochschulen (Edition Stifterverband). Essen.

■ **Dr. Heinrich Seidlmeier**, Professor für Organisation und Wirtschaftsinformatik, Qualitätsbeauftragter der Hochschule Rosenheim, E-Mail: heinrich.seidlmeier@fh-rosenheim.de

**Heinz W. Bachmann:
 Systematische Lehrveranstaltungsbeobachtungen an einer Hochschule
 Verläufe von Lehrveranstaltungen an einer schweizerischen Fachhochschule
 bei Einführung der Bologna-Studiengänge – eine Fallstudie**

Reihe: Hochschulwesen - Wissenschaft und Praxis

Seit Herbst 2006 bieten alle Fachhochschulen der Schweiz Studiengänge organisiert nach dem Bachelor-Master-System an, wie das in der Bologna-Deklaration beschlossen worden war. Einer der Haupttriebfedern des Reformprozesses, neben der akademischen Mobilität und der Vorbereitung der Hochschulabsolventen auf den europäischen Arbeitsmarkt, ist die Steigerung der Anziehungskraft der europäischen Hochschulen zur Verhinderung von brain drain und der Förderung von brain gain. Neben diesem globalen Wettbewerb wird durch die gegenseitige Anrechenbarkeit der Studienleistungen in den verschiedenen Ländern auch die Konkurrenz der Hochschulen untereinander gefördert.

Die Bologna-Reform geht von einem neuen Lehrverständnis aus von der Stoffzentrierung hin zu einer Kompetenzorientierung, begleitet von einem shift from teaching to learning. Der Fokus liegt also nicht beim Lehren, sondern auf der Optimierung von Lernprozessen. Vor dem Hintergrund neuerer Erkenntnisse aus der Lernforschung wird auch deutlich, dass das Vermitteln von Wissen im traditionellen Vorlesungsstil nur noch bedingt Gültigkeit hat. Unter Berücksichtigung der obigen Erkenntnisse müsste man heute eher vom Hochschullernen als von der Hochschullehre sprechen. Die vorliegende Studie wird zum Anlass genommen, ein Instrument vorzustellen, mit dem Lehre systematisch beobachtet werden kann. Mit dem beschriebenen Instrument wird der Frage nachgegangen, inwieweit an der untersuchten schweizerischen Pädagogischen Hochschule die oben beschriebene Neuorientierung in der Lehre schon stattgefunden hat. Mit Hilfe des VOS (VaNTH Observational System) sollen systematisch Lehrveranstaltungsbeobachtungen gemacht und festgehalten werden. Das Ziel dieser Studie ist es, Lehrveranstaltungsverläufe an der untersuchten Pädagogischen Hochschule zu erheben im Hinblick auf die Entwicklung von Kursen in Hochschuldidaktik. Die gefundenen Ergebnisse sollen mit der Schulleitung besprochen werden, vor allem auch auf dem Hintergrund des neuen Lernens an Hochschulen. Basierend auf den gewonnen empirischen Daten und den von der Schulleitung entwickelten Zielen können hochschuldidaktische Kurse geplant und umgesetzt werden. Zusätzlich besteht die Chance, bei einer Wiederholung der Studie in einigen Jahren mögliche Veränderungen in der Lehre festzustellen. Es wird davon ausgegangen, dass das Untersuchungsdesign und die erhobenen Daten nicht nur von Interesse für die betroffene Hochschule sind, sondern generell Fachhochschulen interessieren dürften, die in einem ähnlichen Prozess der Neuorientierung stecken.



ISBN 3-937026-65-7, Bielefeld 2009, 172 Seiten, 24.90 Euro

Bestellung - E-Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

Isolde Kirnbauer & Katharina Stowasser-Bloch

Pilotprojekt „Lehrveranstaltungsbeurteilung mittels Beobachtungsverfahren an der MedUni Wien“



Isolde Kirnbauer



Katharina Stowasser-Bloch

In der sozialwissenschaftlichen Fachliteratur wird die ausschließliche Bewertung der Lehrqualität durch die studentische Einschätzung als zu einseitig eingestuft, da Studierende z.B. die Interaktionsebene zwischen Lehrenden und Studierenden nicht umfassend in ihrer Gesamtheit bewerten können (vgl. Reischmann 2006, S. 144; Bachmann 2009, S.7; Flick 2008, S.25). Hier ist anzuführen, dass die Studierenden vorwiegend mit dem zu Erlernenden beschäftigt sind und/oder im universitären Umfeld akzeptierte Routinen nicht mehr wahrnehmen können sowie über ein zu geringes didaktisches Grundwissen für die Bewertung der Lehrqualität verfügen. In diesem Zusammenhang wird in der sozialwissenschaftlichen Literatur wiederholt die Forderung gestellt, dass die durch Studierende durchgeführte Lehrbewertung in einem Methodenmix (Triangulation¹) aus z.B. Peer Reviews, Selbst-Evaluation, Lehrportfolios, Hörsaalobservationen und Kennzahlen (u.a. Noten bzw. Leistungen der Student/innen) eingebettet sein sollte, um dem Ansatz einer „holistischeren“ Evaluation gerecht zu werden, die mehrere Bewertungsperspektiven berücksichtigt. Diese Überlegungen waren initiativ für die Durchführung eines Pilotprojektes, in welchem die Stabstelle für Evaluation und Qualitätsmanagement an der Medizinischen Universität Wien das alternative Evaluationsverfahren „Beobachtung von Lehrveranstaltungen durch „objektive“ (außenstehende) Beobachter/innen“ für eine zusätzliche Einschätzung der Lehrqualität und als objektiveres Instrumentarium im Rahmen eines zweisejmestri-gen Pilotprojektes erproben wollte.

1. Untersuchungsdesign

Das Beobachtungsprojekt wurde in zwei Phasen umgesetzt. In der **Phase I** (Wintersemester 2007/08) wurden als Beobachtungseinheiten **Vorlesungen**² aus dem ersten bzw. zweiten Studienabschnitt der Undergraduate-Studienrichtung Humanmedizin³ ausgewählt. Hier wurde als Verfahren die nicht-teilnehmende Beobachtung angewendet. Im Rahmen der nicht-teilnehmenden Beobachtung sind Beobachter/innen selbst nicht aktiver Bestandteil des Beobachtungsfeldes. In der **Phase II** (Sommersemester 2008) wurde der Schwerpunkt auf die **Kleingruppenarbeit** im Rahmen von Seminaren⁴ des zweiten Studienabschnitts Humanmedizin gelegt, Methodenwahl war diesmal die teilneh-

mende Beobachtung. Teilnehmende Beobachtung setzt aktive Teilnahme im Untersuchungsfeld voraus. Die zurückhaltende Teilnahme und eine verdeckte Beobachtertätigkeit sind Voraussetzung, um das natürliche Setting nicht zu beeinflussen bzw. zu stören. In dieser Phase des Pilotprojektes lag der Fokus auf der Analyse der Interaktionsebene zwischen Lehrenden und Studierenden. Für das Pilotprojekt wurden Lehrende von den 10% am besten und am schlechtesten bewerteten Lehrenden je Curriculumelement (Modul) des Studienjahrs 2006/07 im Rahmen einer an der MedUni Wien durchgeführten Online-Evaluation ausgewählt und um deren freiwillige Teilnahme gebeten. In der ersten Phase des Pilotprojektes war der Anteil der negativ evaluierten Lehrenden im Vergleich zu dem der positiv evaluierten höher. In der zweiten Phase wurden nur ausschließlich negativ bewertete Lehrende beobachtet.

In beiden Phasen waren zwei bis drei Beobachter/innen in mehreren zeitlich getrennten Einheiten einer Lehrveranstaltung anwesend. Diese ausgewählten Einheiten

¹ Die Bezeichnung Triangulation ist der Mathematik (Geometrie) entnommen und bezieht sich dort auf die geometrische Dreiecksbeziehung. In der Sozialwissenschaft werden – je nach empirischer Erforderlichkeit – verschiedene Erhebungsmethoden miteinander kombiniert. Hierbei beschränkt sich der Methodenmix nicht auf drei Methoden, sondern es können zwei oder mehrere sein.

² Die Veranstaltungsart Vorlesung an der Medizinischen Universität Wien dient bei den Undergraduate - Studienrichtungen der Einführung in die Grundkonzepte und Systematik, dem Aufzeigen des wissenschaftlichen Hintergrundes, der Schaffung von Querverbindungen, der Erklärung von komplizierten Sachverhalten und der Bedeutung für die klinische Anwendung. (Vgl. Mitteilungsblatt der Medizinischen Universität Wien, 17. Stück, Nr. 20, Änderung des Curriculums für das Diplomstudium Humanmedizin, ausgegeben am 30.06.2011) Der Vorlesungsstil ist als frontale Instruktion ohne praktischen Übungsteil für Studierende zu verstehen. Von den Dozent/innen/Lektor/innen wird gewünscht, dass Student/innen während der Vorlesung Fragen stellen bzw. die Student/innen werden von den Vortragenden auch angesprochen, um eine aktive Involvierung der Studierenden in der Veranstaltung zu erzeugen.

³ An der Medizinischen Universität Wien werden drei Undergraduate Studienrichtungen (Diplomstudien Human- und Zahnmedizin und Masterstudium Medizinische Informatik), zwei Doktoratsstudiengänge (PhD und Angewandte Medizinische Wissenschaft) und 13 Universitätslehrgänge angeboten.

⁴ Kleingruppenunterricht an der Medizinischen Universität Wien bedeutet, dass entweder 10 oder 20 Studierenden zu einer Gruppe zusammengefasst werden und in einem eigenen Raum durch einen Lehrenden angeleitet werden. Im Rahmen von Kleingruppen sollen Studierende vor allem die Fähigkeit erlernen, das erworbene Wissen zur Analyse und Lösung von Fragestellungen anzuwenden. Mit dieser Lehrmethode wird vor allem die eigenständige Auseinandersetzung mit theoretischen Problemen auf wissenschaftlicher Basis geschult. (Vgl. ebenda)

wurden von je einer/m Lehrenden abgehalten. Die Beobachter/innen wurden nach bestimmten Kriterien ausgewählt und von der Stabstelle für Evaluation und Qualitätsmanagement eingeschult, siehe dazu Kapitel Auswahl und Vorbereitung der Beobachter/innen (Beobachterschulung). Nach jeder beobachteten Lehrveranstaltungseinheit wurde ein standardisiertes Gespräch mittels Interviewleitfaden mit den jeweiligen Lehrenden über deren Selbstwahrnehmung geführt oder ihnen ein kurzer Fragebogen mit offenen Items zugesendet. Zusätzlich wurde ein standardisierter Beobachtungsbogen von den Beobachter/innen ausgefüllt, der Items mit Ratingskalen (Likertskalen) zu bestimmten Aspekten enthält.

Mittels eines statistischen Tests für die Überprüfung der **Inter-Rater-Reliabilität**⁵ wird vor der Ergebnisdarstellung bzw. -aufbereitung der Beobachtungsdaten festgestellt, welche Bewertungen der zwei bis drei Beobachter/innen je Aspekt bzw. Item (eher) deckungsgleich sind. Bei einer signifikanten Abweichung der Bewertungen muss von einer Ergebnisdarstellung Abstand genommen werden, da man bei Vorliegen von divergierenden Beobachtungsdaten nicht weiß, welche der zwei bis drei Bewertungen je Item (Zahl von der Anzahl der Beobachter/innen abhängig) der Wirklichkeit entspricht. Divergierende Daten können auf subjektive Bewertungen wie auch auf empirisch falsch formulierte Items zurückzuführen sein. Sind die Bewertungspunkte von zwei bis drei Beobachter/innen fast deckungsgleich, so wird aus diesen ein Mittelwert (Median) berechnet.

Im Zuge der Ergebnisaufbereitung bzw. -darstellung werden die einzelnen Bewertungskategorien (Items) zu Oberbegriffen (Aspekten) zusammengefasst. Je nach Aspekt wird ein Index berechnet. Die Indexformel setzt sich additiv aus gleich gewichteten Items zusammen.

In der Phase I fungierten die von der Stabstelle für Evaluation und Qualitätsmanagement ausgewählten und auf neutrale Beobachtung geschulten Student/innen als nicht-teilnehmende Beobachter/innen. Während der Vorlesungseinheiten wurde zusätzlich der Verlauf der Veranstaltung durch die Beobachter/innen mitprotokolliert (wie Dauer von bestimmten Handlungen wie z.B. Diskussionen, Vorträge, Übungseinheiten, ...).

Die an den – für das Pilotprojekt ausgewählten – Lehrveranstaltungen teilnehmenden Studierenden wurden in beiden Phasen nicht über den Einsatz von Beobachter/innen informiert, um das natürliche Setting nicht zu beeinflussen sowie eine künstliche Situation zu erzeugen. Die Beobachtung erfolgte somit für die Studierenden verdeckt.

2. Methodisches Vorgehen

2.1 Konzeption und Entwicklung der Messinstrumente

A) Standardisierter Beobachtungsbogen:

In der Vorphase wurde zuerst festgelegt, welche Beobachtungseinheiten, -gegenstände bzw. Interaktionen (WAS) untersucht bzw. bewertet werden sollen und WIE diese zu beobachten bzw. zu protokollieren sind. In einem weiteren Schritt erfolgte die Bildung von Beobachtungskategorien, die sich auf beobachtbare Inhalte (Beobachtungseinheiten bzw. -gegenstände) beziehen

und zugleich die Variablen mit Merkmalsausprägungen in dem Beobachtungsbogen darstellen. Für die Beurteilung des Ausprägungsgrades je eines beobachteten Verhaltens oder einer Eigenschaft wurde eine Likert-Skala im Beobachtungsbogen zugeordnet (vgl. Bortz/Döring 2002, S. 270, Schnell/Hill/Esler 2005, S.393f). In der Konzeptionsphase wurden didaktisch erfahrene Lehrende und Curriculumdirektor/innen⁶ eingebunden. Als letzter Formalschritt erfolgte die Approbation der Beobachtungskategorien durch den Vizerektor für Lehre.

Laut sozialwissenschaftlicher Fachliteratur wird für die Konzeption des „Beobachtungsbogens“ auf die Einhaltung folgender formaler und inhaltlicher **Qualitätskriterien** hingewiesen (vgl. Schnell/Hill/Esler 2005, S.397):

- Eindimensionalität der Messung,
- Ausschließlichkeit der Kategorien (jedes beobachtete Ereignis nur einer Kategorie zuordenbar),
- Vollständigkeit der Kategorien,
- Beobachtbare Sachverhalte müssen Kategorien zugeordnet sein,
- Begrenzung der Anzahl von Kategorien (aufgrund der begrenzten Wahrnehmungsfähigkeit der Beobachter/innen sollte der Umfang der Kategorien um 10 liegen).

Im Rahmen der Konzeption wurde auf die obig aufgelisteten Kriterien geachtet und im Zuge des Pretests wurde die Einhaltung der Kriterien überprüft, siehe Kapitel Qualitätssicherung.

B) Verlaufsaufzeichnungsbogen (Verlaufsdokumentation): Neben dem Beobachtungsbogen gibt es einen Verlaufsaufzeichnungsbogen⁷, der für die zeitliche Aufzeichnung der Lehrveranstaltungsabläufe bzw. -aktivitäten eingesetzt wird, ohne dass dabei Bewertungen (Ratings) abgegeben werden. Im Rahmen der Pretest-Phase und in der ersten Beobachtungsphase (Beobachtung von Vorlesungseinheiten) fand der Verlaufsaufzeichnungsbogen Verwendung. Ziel des Aufzeichnungsbogens ist die Feststellung, wie lange Handlungen wie Diskussionen, Vorträge und Übungseinheiten dauern. Anhand des Vergleichs der Verlaufsaufzeichnungen aus zwei oder drei Beobachtungen zu derselben Lehrveranstaltungseinheit kann man die Güte der Beobachtung in jedem Einzelfall überprüfen.

⁵ Inter-Rater-Reliabilität bezeichnet das Ausmaß der Übereinstimmung der Einschätzungsergebnisse von zwei oder mehreren BeobachterInnen (Ratern) zum gleichen Zeitpunkt. Zur Bestimmung der Inter-Rater-Reliabilität kann eine Reihe von statistischen Tests (Koeffizienten) eingesetzt werden.

⁶ Ein/e CurriculumdirektorIn ist ein/e Funktionsträger/in an der Medizinischen Universität Wien. Diese/dieser wird aus dem Kreis des wissenschaftlichen Universitätspersonals mit *venia docendi* bestellt. Ein/e CurriculumdirektorIn wird für jede an der MedUni Wien angebotene Studienrichtung bestellt. Ihre/Seine Aufgabe besteht darin, studienrechtliche Bestimmungen gemäß dem österreichischen Universitätsgesetz (UG) 2002 zu vollziehen, insbesondere die Erlassung studienrechtlicher Bescheide in erster Instanz, sowie die Koordination des ihr/ihm zugeordneten Curriculums.

⁷ Beobachtungsschema mit Zeitangabe (wie lange Interaktionen dauern), siehe auch Tafel 29: Beispiel für einen Beobachtungsplan Bortz/Döring 2002, S. 271.

Zur Veranschaulichung:

Verlaufsaufzeichnungsbogen

(Aufzeichnung im 5 Minuten Takt!)

Abbildung 1: Verlaufsaufzeichnungsbogen

(Aufzeichnung im 5 Minuten Takt!)	
10:00 – 10:05	Pkt, Stud_Unpkt, impAnord
10:06 – 10:10	ThV, Hilfsmit (PPT Schärfe), Wa → Erkund, Stud_Bit (PPT), Eingeh
.....
10:21 – 10:25	SachZusatz (weiterf. Lit: 2 min), pBsp.
.....

Codierung:

Z.B.: Pkt. ... Lehrende/r zur veranschlagten Beginnzeit präsent, Stud_Unpkt ... Studierende kommen zu spät, impAnord ... verschafft Ruhe durch verbale Artikulationen, ThV ... theoretischer Vortrag ohne praktische Beispiele, Hilfsmit ... Verwendung von zusätzlichen didaktischen Hilfsmitteln, Wa ... Lehrende/r bemerkt Unruhe bei den Studenten, Erkund ... erfragt den Grund bei den Studierenden, Stud_Bit ... Bitten/Vorschläge in Bezug auf erwünschtes Verhalten mit Erklärung, Eingeh ...Reagieren auf Bitten von Studierenden, SachZusatz ... zusätzliche Sachinformationen für Studierende, pBsp ... Praktische (Fall-) Beispiele vortragen.

Die Aufzeichnung erfolgte entweder mittels vorab festgelegter Code-Wörter oder ad hoc mit selbst getroffener Kurzschreibweise.

C) Interviewleitfaden für die Befragung der Lehrenden: Im Anschluss an Lehrveranstaltungseinheiten, das sind einzelne Veranstaltungstermine, erhalten die jeweiligen Lehrenden die Gelegenheit, Rückmeldung bezüglich des zufriedenstellenden oder erfolgreichen Verlaufs der Lehrveranstaltung bzw. Instruktion oder Moderation der studentischen Gruppe zu machen bzw. die Lehreinheit selbst zu reflektieren. Für diese Befragung wurde ein standardisierter Interviewleitfaden konzipiert.

Im Rahmen der Phase I (Vorlesungsbewertung) wurde direkt nach der Lehrveranstaltung ein Gespräch mit den Lehrenden (über Selbstwahrnehmung) geführt.

In der Phase II (Bewertung der Kleingruppenarbeit) wurden die jeweiligen Lehrenden nicht direkt im Anschluss nach jeder beobachteten Lehrveranstaltungseinheit interviewt, da man vor den teilnehmenden Studierenden nicht die Beobachtungstätigkeit bzw. Beobachterrollen aufdecken wollte. Den Lehrenden wurde daher ein elektronisches Formular mit offenen Fragen zur Lehrveranstaltungseinheit zugesandt. Diese Formulare wurden von den involvierten Lehrenden ausgefüllt und der Stabstelle EQ übermittelt.

2.2 Qualitätssicherung im Rahmen des Beobachtungsverfahrens

Ein Beobachtungsverfahren darf sich nicht auf eine einmalige Momentaufnahme beziehen, sondern soll sich aus der Analyse mehrerer zeitlich getrennter Einheiten einer Lehrveranstaltung zusammensetzen, die durchgehend von einer Person abgehalten werden. Für die Bewertung der Lehrenden müssen zeitgleich in der Lehrveranstaltung zwei bis drei für Beobachtungen geschulte Personen anwesend sein. Neben der ausreichenden Anzahl an Beobachtungseinheiten und der Präsenz von mindestens zwei Beobachtenden gibt es noch weitere Kriterien, auf die hinsichtlich der Qualitätssicherung im Rahmen des Beobachtungsverfahrens geachtet werden sollte. Insbesondere findet die Konzeptionsphase mit Einbeziehung von didaktischen Fachexperten, die Über-

prüfung der Beobachtungsinstrumente im Rahmen eines Pre-Tests, die Schulung der Personen mit Beobachtungsaufgaben, um einen individuellen Bias⁸ zu vermeiden, sowie die methodische/statistische Kontrolle der Beobachtungsleistungen auf Homogenität⁹ und Vollständigkeit in der Aufzeichnung bzw. Bewertung von relevanten Beobachtungsapekten ihre Beachtung.

A) Pre-Testung des Beobachtungsinstrumentes (Beobachtungsbogen, Codierungsplan für die Verlaufsprotokollierung):

Im Rahmen eines probemäßigen Realeinsatzes (Sommersemester '07) wurden die Bewertungskategorien des Beobachtungsbogens auf inhaltliche Validität überprüft. Dafür wurden die Beobachter/innen in ausgewählte Vorlesungen gesetzt, um die Verlaufsprotokollierung zu üben sowie sich mit dem Beobachtungsbogen bzgl. der Anwendbarkeit, Vollständigkeit und eventueller Redundanz der Beobachtungskategorien kritisch auseinanderzusetzen. Zwei bis drei Studierende fungierten als verdeckte, nicht-teilnehmende Beobachter/innen in ausgewählten Lehrveranstaltungen. Diese protokollierten den Ablauf der Lehreinheit mit und füllten nach der Lehreinheit den Beobachtungsbogen aus. Die Beobachtungsergebnisse wurden ausschließlich für das Review der Kategorienschemata verwendet. Die Ergebnisse wurden für eine Beurteilung der jeweiligen Lehrenden nicht verwendet. Die inhaltliche Validität der einzelnen Beobachtungssitems wurde mittels der Verlaufsaufzeichnungsprotokollierung und einer Feedbackrunde (Diskussion) mit Beobachter/innen bzgl. der Kriterien Messbar-, Beobachtbarkeit und Vollständigkeit der Kategorien analysiert. Die Begrenzung der Anzahl der Kategorien bzw. Items auf zehn wurde bewusst in der Konzeptionsphase aufgeweicht bzw. übertroffen. Die Beobachter/innen wurden in der Feedbackrunde nach der Wahrnehmungsgrenze für die Anzahl der Kategorien befragt. Hierbei stellte sich heraus, dass die Kategorienanzahl im Beobachtungsbogen nicht die Erschwernis in der Beobachtertätigkeit bedingte, sondern das Merken von zahlreichen Codes (Word-Kürzel) für die Verlaufsaufzeichnung. Aufgrund dieses Feedbacks wurde die Verwendung von Codes nicht mehr strikt gefordert, sondern es wurde die Verwendung von eigenen Kürzeln empfohlen. Diese Kürzel mussten dann anschließend in die vorgegebenen Codes übersetzt werden, sodass die Abkürzungen wieder für andere verständlich wurden. In der Feedbackrunde wurden zusätz-

⁸ Der Begriff Bias steht in der Empirie für eine systematische Antwortverzerrung. Ein Bias führt auf einen Fehler in der Erhebung zurück und besagt, dass nicht der wahre Wert (=tatsächliche Situation) abgebildet wird. Die Ursache für eine systematische Verzerrung kann bei den Fragen (Items), Befragten, Interviewer/innen bzw. Beobachter/innen liegen. Die Verzerrung kann ab drei Beobachter/innen festgestellt werden. Bei zwei Beobachter/innen, die (stark) abweichende Bewertungen zu einem Aspekt abgegeben haben, weiß man nicht, welche Beobachtungsangabe die korrekte ist, somit muss die Ergebnisdarstellung für die betreffende Kategorie (Eigenschaft) unterlassen werden.

⁹ Im Rahmen der Überprüfung auf Homogenität wird auf gleiche Antworttendenzen zwischen ausgefüllten Bewertungsbogen und Verlaufsprotokollierung je Beobachter/innen (Intra-Reliabilität) und zwischen Beobachter/innen (Inter-Reliabilität) geachtet.

liche Beobachtungskategorien/-aspekte durch die Beobachtenden aufgeworfen, die im Beobachtungsbogen nicht enthalten, aber für die Bewertung essentiell waren. Diese neuen Kriterien wurden in der Feedbackrunde nach dem Prinzip der Dialektik diskutiert und anschließend entweder abgelehnt oder aufgrund breiter Zustimmung in den Beobachtungsbogen bzw. in den Codierungsplan für die Verlaufsprotokollierung aufgenommen. Bei Feststellung von Redundanz (siehe Kriterium „Ausschließlichkeit der Kategorien“ im vorausgehenden Kapitel „Konzeption und Entwicklung von Messinstrumenten“) von Kategorien wurden die betreffenden aus dem Beobachtungsbogen gestrichen bzw. nicht aufgenommen.

B) Auswahl und Vorbereitung der Beobachtenden (Beobachtungsschulung):

Die auszuwählenden Beobachter/innen mussten für die erste Pilotphase als Anforderungsprofil ein sozialwissenschaftliches Studium im fortgeschrittenen Semester sowie praktische Erfahrung im Bereich der Empirie aufweisen. Das Anforderungsprofil für die zweite Phase verlangte aufgrund des teilnehmenden Beobachtungsverfahrens¹⁰ nach Studierenden der Medizin aus höheren Semestern. Als Auswahlkriterium fungierte das schriftliche Qualifikationsprofil (Lebenslauf) der Bewerberin bzw. des Bewerbers sowie das persönliche Bewerbungsgespräch.

Auf die Vorbereitung der Beobachter/innen auf den Realeinsatz wurde im Rahmen der Vorphase (Pre-Test) größter Wert gelegt, da das Hauptproblem bei der Beobachtungsmethode die beobachtenden Personen an sich sein können (Bias-Anfälligkeit, siehe Erläuterung in Fußnote unter Punkt A), wenn zu wenig auf die Beobachtungstätigkeiten vorbereitet bzw. sensibilisiert oder wenn das Beobachtungsverfahren zu gering strukturiert bzw. standardisiert angelegt wurde. Zu der Beobachtungsschulung zählt eine Einführung in das Konzept und die Absichten (Ziele) der gesamten Untersuchung sowie eine Sensibilisierung gegenüber relevanten Verhaltens-, Handlungs- und Interaktionsformen (vgl. Bortz/Döring 2002, S. 273). Die vorab hergestellte Eini-gung (gemeinsames Verständnis) über die Bedeutung von Indikatoren und Kategorien unter den eingesetzten Beobachter/innen ist essentiell. Wenn die Eindeutigkeit von Beobachtungsindikatoren und Codierungen nicht gegeben ist, dann kann die Messung nicht den Standard der Validität, Reliabilität und Objektivität erfüllen. Im Rahmen eines Pre-Tests kann mittels der Verlaufsprotokollierung durch Beobachter/innen nachvollzogen werden, wie genau Beobachter/innen arbeiten bzw. wer Wichtiges von Unwichtigen trennen kann. Wird hierbei eine schlechte Qualität in der Beobachtungstätigkeit festgestellt, muss die beobachtende Person ausgewechselt werden.

Folgende Beobachtertätigkeiten laufen fast zeitgleich ab und können deshalb Fehler mit sich bringen (vgl. Schnell/Hill/Esser 2005, S. 400 ff):

- Beobachter/in „entdecken“ Daten/Ereignisse (Acts),
- zeichnen Daten auf (Protokollierung) und
- „verarbeiten“ Daten beurteilend (Beurteilung und Codierung).

Die Verarbeitungsphase ist besonders anfällig für Beobachtungsfehler, da hier Urteile gebildet und Schlussfolgerungen gezogen werden müssen und folglich können extreme Ereignisse zu milde beurteilt werden oder ein weiterer Urteilsfehler wie der Halo-Effekt (Verzerrung von Urteilen aufgrund eines besonderen Merkmals einer Person bzw. eines Ereignisses, die für alle nachfolgenden Urteile Nachwirkungen haben) kann auftreten. Instruktionen dürfen sich deshalb nicht nur auf die Schulungsphase beziehen, sondern während der Feldphase muss eine kontinuierliche Kontrolle und Überwachung der Beobachter stattfinden!

C) Methodische Kontrolle der Beobachtungsleistungen:

Im Unterschied zur Teilnehmer/innen- und Student/innenbewertung ist eine bestimmte Anzahl an Respondent/innen für die Aussagekraft der Ergebnisse bzw. für die Güte der Repräsentativität nicht erforderlich, da generell eine Antworthomogenität pro Item (Bewertungsaspekt) unter den Beobachter/innen vorliegen muss, um die Bewertungsergebnisse weiter aufbereiten bzw. als gültige Aussage verwenden zu können. Im Vorfeld zur Aufbereitung der Ergebnisse muss somit eine statistische Überprüfung der Übereinstimmung der Beobachtungsdaten bei Einsatz von zwei oder drei Beobachter/innen durchgeführt werden. In diesem Pilotversuch wurden die Beobachtungsdaten der Beobachter/innen mittels eines statistischen Tests für die Feststellung der „Inter-Rater-Reliabilität“ überprüft.

Für die Überprüfung der „**Inter-Rater-Reliabilität**“ können verschiedene Korrelationskoeffizienten verwendet werden, wie z.B. der Spearman'sche¹¹, das Cohens Kappa¹² oder der Kendalls Tau-b¹³ Korrelationskoeffizient. Diese Koeffizienten stellen den Grad der Übereinstimmungen der Beurteilungen durch Beobachter/innen, die dasselbe Objekt zu bewerten hatten. In dieser Abhandlung wurde für die Überprüfung der Übereinstimmung der Beobachtungsdaten das Cohens Kappa sowie der Spearman'sche Korrelationskoeffizient verwendet.

Bei Einsatz von drei Beobachter/innen kann ggfls. die Person (Ausreißer) ermittelt werden, die im Vergleich zu den anderen Beobachter/innen mit ihren Beobachtungswerten abweicht bzw. auffällt. Diese abweichende Bewertung wird anschließend eliminiert und fließt nicht in die Ergebnisdarstellung ein. Bei Feststellung von abweichenden Bewertungsdaten zweier Beobachter/innen bei einem Bewertungsaspekt muss eine Ergebnisdarstellung unterbleiben (siehe dazu Definition ad Bias in der Fußnote unter Punkt A).

¹⁰ Teilnehmende Beobachtung setzt am Unterricht mitwirkende Tätigkeit (aktive Mitarbeit) neben der Beobachtertätigkeit voraus.

¹¹ Spearman'sche Korrelationskoeffizient ist ein Maß für die Feststellung der Übereinstimmung zwischen Beobachter/innen und Urteiler/innen (Ratern), das für metrisch- oder ordinalskalierte Merkmale geeignet ist und Werte von -1 bis + 1 annehmen kann.

¹² Cohens Kappa ist ein Maß für die Feststellung der Übereinstimmung zwischen Beobachter/innen und Urteiler/innen (Ratern), das für nominal- oder ordinalskalierte Merkmale geeignet ist und Werte kleiner, gleich Null oder bis zu + 1 annehmen kann.

¹³ Kendall Tau-b ist ein Zusammenhangsmaß, das vorrangig für ordinalskalierte Daten verwendet wird und Werte zwischen -1 und +1 annehmen kann.

3. Durchführung und Ergebnisse des Beobachtungsverfahrens

3.1 Nicht-teilnehmende Beobachtung im Rahmen von Vorlesungseinheiten – Phase I

Im Rahmen der Beobachtungsphase I wurden Student/innen der Sozialwissenschaften, die aufgrund ihrer Ausbildung im Rahmen des Studiums ausreichend Erfahrung mit empirischen Beobachtungsverfahren sammeln konnten, als neutrale Beobachter/innen im Rahmen einer nicht-teilnehmenden Beobachtung von Vorlesungen eingesetzt. Diese Beobachter/innen wurden – wie schon unter dem Abschnitt Qualitätssicherung erwähnt – von der Stabstelle für Evaluation und Qualitätsmanagement in ihre Beobachterrolle eingeschult. Das Beobachtungskonzept sieht eine permanente Verlaufsprotokollierung während der Vorlesung, ein Interview¹⁴ mit den jeweiligen Lehrenden über deren Selbstwahrnehmung nach der Veranstaltung und eine abschließende Beobachtungsbewertung mittels eines für die Vorlesung speziell entworfenen Beobachtungsbogens mit Ratingskalen vor. Am Ende der Beobachtungsphase I fand eine Feedbackrunde mit den Beobachter/innen und den Mitarbeiter/innen der Stabstelle für Evaluation und Qualitätsmanagement statt, um die einzelnen Eindrücke über den Verlauf der Beobachtung zu sammeln und die Bewertungen durch die Beobachter/innen den Selbstbewertungen der betreffenden Lehrenden diskursiv gegenüber stellen zu können.

Als Beobachtungsgegenstand wurden Vorlesungseinheiten von Curriculumelementen (Modulen) des ersten und zweiten Studienabschnitts der Studienrichtung Humanmedizin an der MedUni Wien ausgewählt. Untersucht wurden vorwiegend Lehrveranstaltungen von Lehrenden, die im Rahmen der Online-Evaluation als „signifikant“ schlecht oder sehr gut durch die Studierenden eingestuft wurden.

Im Beobachtungsbogen sind fünf Bewertungsebenen (Hauptaspekte) für die Bewertung der jeweiligen Vortragenden vorgesehen, wie didaktische Aufbereitung/Vermittlung des Lernstoffes¹⁵, Sprechqualität¹⁶, qualitativer Einsatz von Lehr-Medien (wie insbes. Powerpoint Präsentationsfolien)¹⁷, soziale Kompetenz (Einbindung der Studierenden, Stimmungsvermittlung, Umgang mit Unruhe) und Zeitmanagement¹⁸.

Für die Bewertung des Verhaltens der Studierenden im Rahmen einer Vorlesung wurden Aspekte wie Pünktlichkeit, Engagement und Interesse¹⁹ ebenfalls berücksichtigt.

Die Beobachtungsdaten wurden einem statistischen Test für die Feststellung der „Inter-Rater-Reliabilität“ unterzogen, um zu testen, ob die Bewertungen durch die Beobachter/innen (annähernd) deckungsgleich sind. Für die Überprüfung der Inter-Rater-Reliabilität wurde der Cohens Kappa Koeffizient verwendet. Generell konnte eine hohe Deckungsgleichheit zwischen den Bewertungen der zwei bis drei Beobachter/innen festgestellt werden. Wie bereits in der Einleitung beschrieben, wurde ein Index je Aspekt im Zuge der Ergebnisaufbereitung berechnet. Die Teilaspekte wurden gleich gewichtet in die Indexformel aufgenommen.

Anhand zahlreicher Bewertungskriterien im Rahmen des

Beobachtungsverfahrens wurden die vom Studienjahr 2006/07 negativ evaluierten Lehrenden im Durchschnitt gut bewertet, mit Ausnahme des Aspektes der sozialen Kompetenz. Die Lehrenden animierten laut Angaben der Beobachter/innen die Studierenden nur teilweise dazu, sich in die Veranstaltung einzubringen.

Die Studierenden wurden ebenso wie die jeweiligen Lehrenden mittels des Beobachtungsverfahrens beurteilt. Im Beobachtungsbogen wurde unter anderem die Störung der Veranstaltung durch studentische Unruhe erhoben. Laut den Angaben der Beobachter/innen im Fragebogen sowie aus der Verlaufsprotokollierung zu entnehmen, gab es ab und zu Störungen seitens der Studierenden in den hinteren Bankreihen der Hörsäle.

Nach jeder beobachteten Vorlesungseinheit wurden die Lehrenden von den teilnehmenden Beobachter/innen interviewt. Die Ergebnisse aus der Befragung ergaben, dass es den Lehrenden laut der Selbsteinschätzung gelungen ist, die Lerninhalte so zu vermitteln, wie sie es beabsichtigt hatten. Die mit den Lehrenden geführten Interviews lassen keine negative Beeinflussung der Veranstaltung durch Studierende erkennen.

3.2 Teilnehmende Beobachtung im Rahmen von Kleingruppen (POL/FBL)

Für die Beobachtertätigkeit in Kleingruppen im Rahmen der Line-Elemente²⁰ ProblemOrientiertes Lernen (POL)²¹ und FallBasiertes Lernen (FBL)²² wurden zwei Medizinstudent/innen höheren Semesters eingeschult. Weil die teilnehmende Beobachtung Mitarbeit in der Gruppe bedingt, kamen Student/innen der Sozialwis-

¹⁴ Dabei wurden die beiden Fragen "Ist es Ihnen gelungen, Lerninhalte so zu vermitteln, wie sie es beabsichtigt haben?" und "Gab es aus ihrer Sicht (positive/negative) Aspekte, die den Erfolg der Lehrveranstaltung beeinflusst haben?" gestellt.

¹⁵ Darunter fallen die Teilaspekte wie "Vermittlung eines Überblicks über Lehrziele am Beginn der Veranstaltungseinheit", "Inhalte gut strukturiert", "anlassbezogener Einsatz von Übungsbeispielen", "verständnisförderlicher Einsatz von visuellen Darstellungen (Abbildungen/Schaubildern)", "wirkte flexibel in Vortragsgestaltung", "erklärte verständlich", "wirkte vorbereitet", "bei Folieneinsatz freies Sprechen, kein Ablesen".

¹⁶ Teilaspekte wie "deutliche Sprechweise/Sprache", "gut hörbar bis in die letzte Reihe", "Mitverfolgbarkeit des Vortrages war aufgrund der Redegeschwindigkeit immer gewährt".

¹⁷ Teilaspekte wie „Einsatzhäufigkeit während der Lehrveranstaltungseinheit“, „Qualität der Aufbereitung der Folien“: gut strukturiert, stichwortartig, aber verständlich usw. und „Optimale Geschwindigkeit in der Folienabfolge“.

¹⁸ Teilaspekte wie "Lehrende/r zur öffentlichen Beginnzeit präsent" und "Lehrende/r wirkte koordiniert/gut organisiert (teilt die Veranstaltungszeit sinnvoll in Vortrag, Diskussion, Klärung von Fragen) ein".

¹⁹ Teilaspekte wie "Studierende zur veröffentlichten Beginnzeit präsent", "stellten oft Fragen/leisteten aktive Mitarbeit" und "wirkten aufmerksam und am Thema interessiert".

²⁰ Im Rahmen von Line-Elementen wird der Bezug zur klinisch-praktischen Tätigkeit hergestellt und entsprechende klinische Fähigkeiten und Fertigkeiten werden erlernt. Ein Line-Element dauert ein ganzes Semester.

²¹ Ausgangspunkt für die Lehrveranstaltung bzw. Lehrmethode POL ist eine Problemstellung oder ein Fallbeispiel aus der Berufspraxis. Diese Fallbeispiele sollen die Studierenden dazu aktivieren, die von den Fallautor/innen intendierten Lehrziele zu verfolgen und vorhandene Wissenslücken zu schließen. Die Kleingruppenarbeit findet unter Anleitung einzelner Tutor/innen statt. Die Tutor/innen übernehmen in manchen Varianten des POL sowohl die Leitung der Lehrveranstaltung, als auch deren Moderation. Die Tutor/innen sollen hier nicht fachlich eingreifen und zwischen falsch und richtig werten, dennoch sollen sie – auf den Fall vorbereitet – die Studierenden im Rahmen der Problemlösung lenken können (vgl. Mitteilungsblatt der Medizinischen Universität Wien, 17. Stück, Nr. 20, Änderung des Curriculums für das Diplomstudium Humanmedizin, ausgeben am 30.06.2011).

senschaften aufgrund dieses Methodensettings nicht in Frage. Als Beobachtungseinheiten wurden die jeweils ersten vier Lehrveranstaltungs-Einheiten von POL und FBL gewählt, da der anfängliche Einsatz von „fremden Teilnehmer/innen“ bei den Studierenden aufgrund der noch nicht entstandenen Gruppenbildung unter den Studierenden, die die LV besuchten, nicht auffallend ist bzw. nicht als verdächtig hinterfragt wird. Die Lehrenden wurden über die Teilnahme der Beobachter/innen an der Lehrveranstaltung informiert. Die Studierenden wurden nicht eingeweiht, um keine „künstliche“ bzw. verstellte Situation seitens der Studierenden entstehen zu lassen. Als Beobachtungseinheiten wurden POL/FBL-Gruppen von Lehrenden gewählt, die im vorausgehenden Studienjahr 2006/07 im Rahmen der Modul-Online-Evaluation schlechte Bewertungen durch Studierende erzielt hatten.

Die Beobachtungseinheiten wurden jeweils nach Ende jeder Lehrveranstaltungseinheit an Hand eines Beobachtungsbogens von je einer BeobachterIn bewertet.

Auf der Bewertungsebene der jeweiligen Lehrenden lassen sich folgende Hauptaspekte konstatieren, wie didaktische Vorgehensweise²³, Diskussions- bzw. Gruppenmoderation²⁴, Sprechqualität²⁵, soziale Kompetenz²⁶ und Pünktlichkeit.

Die Studierenden wurden nach dem Aspekt Pünktlichkeit und Interesse bzw. aktive Mitarbeit²⁷ bewertet. Wenn einzelne Studierende die Moderation übernahmen – wie im Rahmen der Lehrveranstaltung POL vorgesehen –, dann wurden diese ebenso nach der Art und Qualität der Moderation (Teilaspekte wie bei Lehrenden, siehe dort unter Fußnote) bewertet.

Die Beobachtungsdaten wurden erneut dem statistischen Test für die Feststellung der „Inter-Rater-Reliabilität“ mittels des Cohens Kappa Koeffizienten unterzogen, um zu analysieren, ob die Bewertungen durch die Beobachter/innen (annähernd) deckungsgleich sind. Generell konnte eine hohe Deckungsgleichheit zwischen den Bewertungen der zwei Beobachter/innen festgestellt werden. Wenn marginal abweichende Beobachtungsdaten bei den einzelnen Aspekten (Items) konstatiert wurden, dann wurden die Beobachtungsdaten gemittelt, d.h. Medianwerte wurden bestimmt bzw. bei Vorliegen von nur zwei Beobachtungsdaten ein Mittelwert berechnet²⁸. Bei größerer Divergenz der Beobachtungsdaten pro Item wurde die Mittelwertberechnung unterlassen. Ebenso wurde – wie bereits im Einleitungsteil beschrieben – ein Index je Aspekt für die Ergebnisaufbereitung berechnet. Die Teilaspekte wurden gleich gewichtet in die Indexformel aufgenommen. Als Index wurde das arithmetische Mittel gewählt.

Die beobachteten Lehrenden, die im Studienjahr 2006/07 im Rahmen der Modul-Online-Evaluation bzgl. ihres Kleingruppenunterrichts durch Studierende negativ bewertet wurden, erzielten bei den für das Beobachtungsverfahren ausgewählten Lehrveranstaltungseinheiten des SS'08 hingegen durchgehend positive Bewertungen.

Die Lehrenden erhielten die Möglichkeit, ein schriftliches Feedback nach jeder beobachteten Lehrveranstaltungseinheit abzugeben. Dafür wurden Feedbackbögen zum elektronischen Ausfüllen an die Lehrenden ver-

schickt, d.h. die Beobachter/innen mussten keine Interviews – wie in Beobachtungsphase I – am Ende der LV-Einheit führen, weil dies zu auffallend gewesen wäre. Ausgefüllte Feedbackbögen wurden seitens der Lehrenden direkt an die Stabstelle EQ übermittelt. Wie in der Beobachtungsphase I wurden offene Fragen gestellt, die sich vom Inhalt ähneln²⁹.

Laut dem von den beobachteten Lehrenden selbst abgegebenen Feedback stufen sie ihre eigene Moderation der Gruppe im Großen und Ganzen als gut gelungen ein. Es gab marginale Störungen seitens der Studierenden, aber überwiegend wirkte die Gruppe interessiert und arbeitete aktiv mit. Ebenso wurde der Moderationseinsatz der Studierenden als erfolgreich und gelungen bezeichnet.

Am Ende der Beobachtungsphase II fand eine Feedbackrunde statt. Daran nahmen die bewerteten Lehrenden der Kleingruppenarbeit, die zwei Beobachter/innen und zwei Personen der Stabstelle Evaluation und Qualitätsmanagement (Moderatorin und Protokollführerin) teil. In der Feedbackrunde ging es um den Austausch der Eindrücke von Beobachter/innen und beobachteten Lehrenden bezüglich funktionierender Interaktionsebene zwischen Lehrendem und Studierenden. Der Moderation der Diskussionsrunde lag ein Gesprächsleitfaden zugrunde, der folgende Fragen beinhaltet: (1) Wie ist die konstruktive Mitarbeit bzw. das Interesse der FBL- und POL-Gruppe einzustufen?; (2) Wenn es negative Erfahrungen mit der Gruppe gab, welche Ursachen waren bzw. könnten dafür verantwortlich gewesen sein?; (3) Welche Lösungsansätze gäbe es für die Behebung der Problematik (Spannung/Missverständnisse zwischen Lehrendem und Studierenden)? und (4) Pro & Contra des Beobachtungsverfahrens für Veranstaltungseinheiten durch objektive, geschulte Beobachter/innen.

²² Ziel von FBL ist es, den Stoff aus den vorangegangenen Vorlesungen zu vertiefen und anzuwenden. Medizinische Fälle werden durch die jeweiligen Lehrenden vorgestellt. Die Studierenden werden durch die einzelnen Lehrenden, die als Moderator/innen fungieren, zur Diskussion angeregt. Die Studierenden werden auf klinisches Denken und Arbeiten vorbereitet. (Vgl. ebenda)

²³ Darunter fallen Teilaspekte wie "Vermittlung eines Überblicks über Lernziele bzw. Inhalte der Lehrveranstaltung in der 1. Einheit", "Austeilung eines Fallbeispiels am Beginn der LV-Einheit", "Verständliche Erklärungen", "Gute Vorbereitung", "Erfragung der Verständlichkeit des Lernstoffes".

²⁴ Teilaspekte wie "Gute Moderation der Diskussion zu den Fällen", "Gute Unterstützung der Gruppe, eigene Ideen und Lösungsansätze zu entwickeln", "Deutliche Sprechweise" und "Verständliche Erklärungen".

²⁵ „Deutliche Sprechweise/Sprache“.

²⁶ Teilaspekte wie "Einbindung der Studierenden in Diskussion", "Stimmungsvermittlung" und "Umgang mit Unruhe".

²⁷ Teilaspekte wie "Studierende zur veröffentlichten Beginnzeit präsent", "Studierende verließen vor Lehrveranstaltungsende den Unterrichtsraum", "Studierende wirkten aufmerksam und am Thema interessiert", "Die meisten Teilnehmer/innen beteiligten sich aktiv an der Gruppendiskussion bzw. Falllösung", "Zusammenarbeit der Teilnehmer/innen war konstruktiv".

²⁸ Aus zwei bis drei Beobachtungsdaten je Item wurde der Medianwert berechnet, da die Ratingskalen im Beobachtungsbogen durchgehend ordinal sind. Ein Medianwert aus zwei Beobachtungsdaten – aufgrund zweier Beobachter/innen – stellt immer zugleich einen arithmetischen Mittelwert dar.

²⁹ Fragen wie (1) "Ist Ihnen die Moderation der Gruppe leicht gefallen?" Oder wenn Studierende/r ModeratorIn war: "Ist dem Studierenden die Moderation gelungen?" und (2) "Gab es aus Ihrer Sicht [a] positive, b) negative] Aspekte, die den Prozess der Lösungsfindung in der Gruppe (bzw. den Unterrichtsertrag) beeinflusst haben?"

Die Feedbackrunde ergab, dass zwischen Lehrenden und Beobachter/innen zwei unterschiedliche Ansichten vorherrschend sind. Die beiden Lehrenden bekunden durchgehend interessierte Studierende. Die beiden Beobachter/innen haben einige Studierende als störend und desinteressiert erlebt und meinen, dass sie eine studentische Gruppendynamik in Richtung negativem Verhalten gegenüber einer bestimmten Lehrperson vernommen hätten.

Die Ursache für desinteressierte Studierende wird von den Lehrenden auf die Form der Veranstaltung bzw. auf fehlende Voraussetzungen der Studierenden, wie Teamfähigkeit, zurückgeführt. Ein Druckmittel einzusetzen wäre laut den Lehrenden in diesen Fällen nicht effizient, weil Desinteresse nur durch Motivation beseitigt werden kann.

Seitens der Lehrenden wird eine Anlaufstelle gefordert, an die man sich wenden kann, wenn die Situation mit der Gruppe verfahren ist und es keine weiteren Möglichkeiten zur Lösung des Konflikts gibt. Eine Möglichkeit zur Supervision der Gruppe sollte laut der beiden Lehrenden angeboten werden, wenn sich die Gruppendynamik zum Negativen für die jeweiligen Lehrenden entwickelt.

4. Resümee

Das Beobachtungsverfahren für die Bewertung von ausgewählten Lehrveranstaltungen wurde als Pilotversuch an der MedUni Wien im Studienjahr 2007/08 durchgeführt. Mit der Durchführung des Pilotprojekts „Lehrveranstaltungsevaluation mittels Beobachtungsverfahren“ an der MedUni Wien wurden folgende Ziele bzw. Intentionen verfolgt:

- 1.) Die Ergänzung der studentischen Lehrendenbewertung durch eine objektivere Beobachtung der Interaktionsebene zwischen den jeweiligen Lehrenden und den Studierenden,
- 2.) den durch die herkömmliche Lehrveranstaltungsevaluation schlecht bewerteten Lehrenden die Möglichkeit einzuräumen, sich alternativ durch das Beobachtungsverfahren bewerten zu lassen und
- 3.) das Beobachtungsverfahren als objektivierbares Instrumentarium zu testen.

Das Pilotprojekt gliederte sich in zwei Phasen. In der Phase I lag der Fokus auf der Beobachtung von Vorlesungen durch geschulte Beobachter/innen (Studierende der Sozialwissenschaften) und Phase II – Beobachtung von POL/FBL – Kleingruppen durch geschulte Studierende höheren Semesters der Humanmedizin. Im Rahmen des Beobachtungsprojektes wurden drei verschiedene Messinstrumente eingesetzt: a) Beobachtungsbogen mit Ratingskalen, b) Verlaufsaufzeichnungsbogen für die zeitliche Registrierung von Handlungs-/Ereignisabfolgen (ausschließlich im Rahmen von Vorlesungen) und c) Interviews mit beobachteten Lehrenden am Ende jeder Lehrveranstaltungseinheit.

Die von dem Beobachtungsverfahren betroffenen Lehrenden wurden vorab um Erlaubnis gefragt bzw. um ihre freiwillige Teilnahme gebeten. Die Studierenden wussten in beiden Phasen nichts von dem Einsatz von Beobach-

ter/innen, um das natürliche Setting nicht zu beeinflussen. In der Phase II wurde das Messinstrument Verlaufsprotokollierung aufgrund der zu befürchtenden Auffälligkeit aus dem Methodensetting herausgenommen, um keine künstlich erzeugte Situation unter den Studierenden zu bewirken.

Das Interview mit den Lehrenden jeweils am Ende jeder beobachteten Lehrveranstaltungseinheit diente dazu, der Fremd- eine Selbstbewertung gegenüberzustellen. Durch diesen Vergleich kam deutlich zum Vorschein, dass diese Einschätzungs- bzw. Bewertungsdaten bedingt durch den unterschiedlichen Blickwinkel divergieren können, z.B., dass einige beobachtete Lehrende die Unruhe der Studierenden im Raum gar nicht wahrgenommen haben.

Das Pilotprojekt „Beobachtung von Lehrveranstaltungen“ konnte ohne unvorhersehbare Beeinträchtigungen durchgeführt werden.

Die in den Lehrveranstaltungseinheiten eingesetzten Beobachter/innen wichen mit ihren Bewertungen bei einzelnen Aspekten nur marginal voneinander ab. Die Beobachtungsdaten konnten somit für die Ergebnisaufbereitung verwendet werden. Das Beobachtungsverfahren mit Einsatz von geschulten Beobachter/innen ist aufgrund der positiven Ergebnisse aus der Überprüfung der Homogenität der Beobachtungsdaten (Inter-Rater-Reliabilität) als objektiv einzustufen.

Hervorzuheben ist, dass die vom Studienjahr 2006/07 negativ bewerteten Lehrenden durch das Beobachtungsverfahren in den einzelnen Bewertungskriterien gut abschnitten, mit Ausnahme der sozialen Kompetenz im Rahmen der Phase I (Vorlesungen).

5. Ausblick

Da sich das Beobachtungsverfahren in dem Pilotversuch als ein zusätzliches, objektives Erhebungsinstrument bewährt hat, wird der Einsatz des Verfahrens mittels höhersemestrigen – für das Beobachtungsverfahren geschulten – Studierenden insbesondere dann empfohlen, wenn die Interaktion zwischen den jeweiligen Lehrenden und Studierenden nicht (mehr) funktioniert, um die Ursachen zu eruieren. Wird weiterhin festgestellt, dass die/der Lehrende mit einer unruhigen bzw. der Mitarbeit verweigernden Studierenden-Gruppe zu tun hat, sollte eine Feedback-Runde formiert werden. Die teilnehmenden Beobachter/innen sollen hierbei als Moderator/innen der Feedbackrunde fungieren. In dieser Feedbackrunde kommt die SOFT-Methode zum Einsatz. Das Akronym SOFT steht für Satisfaction, Opportunities, Faults und Threats. Diese Aspekte sollten in dieser genannten Reihenfolge mit der Gruppe und den jeweiligen Lehrenden abgehandelt werden. Zu jedem Aspekt gibt es spezifische Fragen, die in der nachfolgenden Tabelle ersichtlich sind. Die einzelnen Meldungen werden von der beobachtenden Person protokolliert und zuletzt der Gruppe als Vorlage für die gemeinsame Lösungsfindung zur Verfügung gestellt.

Nachdem mittels der SOFT-Methode geklärt ist, wie die Gegenwart (IST-Zustand) erlebt wird, an welchen Anhaltspunkten die Stärken und Schwächen erkennbar sind und welche Chancen und Risiken für die Zukunft

(SOLL-Zustand) gesehen werden, lassen sich konkrete Handlungsschritte ableiten. Diese werden folglich in einem Aktionsplan festgehalten. Der Aktionsplan unterstützt eine gezielte Qualitätsentwicklung für die Lehrveranstaltung bzw. ein kooperatives Miteinander.

Abbildung 2: Mögliche Impulse

Gegenwart	Zukunft
<p>Satisfaction (Zufriedenheit): An welchen Anhaltspunkten lassen sich Stärken erkennen?</p>	<p>Opportunities (Möglichkeiten, Chancen): Hier liegen unsere Chancen. Da haben wir Entwicklungsmöglichkeiten.</p>
<p>Faults (Fehler, Schwächen, Probleme): Das läuft (noch) nicht so gut. An welchen Anhaltspunkten lassen sich Verbesserungsbereiche erkennen?</p>	<p>Threats (potentielle Gefahren) Worauf müssen wir achten?</p>

tischen und interaktiven Einheiten) kann die Veranstaltungsgestaltung (Lehrmethode bzw. Lehrstil) im Kontext der vorgegebenen Lehrziele (genannt „Lernzieldatenbank“) genauestens beleuchtet werden. Die Hospitation durch Fachkolleg/innen mit anschließender Bewertung durch diese Peers wird (nach der Fachliteratur) zumeist bei Junglehrenden angewandt, um jenen ein konstruktives Feedback durch erfahrene Lehrende zuteil werden zu lassen.

Literaturverzeichnis:

Bachmann; H. W. (2009): Systematische Lehrveranstaltungsbeobachtungen an einer Hochschule. Bielefeld.
 Bortz, J./Döring N. (2002): Forschungsmethoden und Evaluation. Berlin, Heidelberg.
 Flick, U. (2008): Triangulation. Eine Einführung. Wiesbaden.
 Reischmann, J. (2006): Weiterbildungsevaluation. Lernerfolge messbar machen. Augsburg.
 Schnell, R./Hill, P./Esler, E. (2005): Methoden der empirischen Sozialforschung, München.

Vor der Feedbackrunde kann die Moderation von Kleingruppen auch an Student/innen, die im Unterricht als störend und/oder desinteressiert erlebt wurden, abgegeben werden, um die Probleme der Gruppe herauszufinden. Nach der Feedbackrunde unter Einsatz der SOFT-Methode wird an der Umsetzung des Aktionsplans gearbeitet. Die jeweilige Lehrperson übernimmt wieder die Moderation bzw. Instruktion der Studierenden.

Anstatt der geschulten Beobachter/innen können ebenso Fachkolleg/innen (Peers) die Beobachtungsrolle übernehmen. Von einer verdeckten Beobachtungsrolle für die an der Veranstaltung teilnehmenden Studierenden kann hierbei nicht mehr ausgegangen werden. Wenn man als Universität die Art und das Ausmaß der Umsetzung der vorgegebenen Lehrziele (genannt „Lernzieldatenbank“) überprüfen möchte, dann ist hingegen ein Verfahren mittels Peers unabdingbar. Mittels zusätzlich permanenter Aufzeichnung des Verlaufs der Veranstaltungseinheit (zeitliches Ausmaß der theoretischen/prak-

■ **Dr. Isolde Lydia Kirnbauer**, Stabstelle für Evaluation und Qualitätsmanagement der Medizinischen Universität Wien,
 E-Mail: isolde.kirnbauer@meduniwien.ac.at
 ■ **Dr. Katharina Stowasser-Bloch**, Leiterin der Stabstelle "Evaluation und Qualitätsmanagement" der Medizinischen Universität Wien,
 E-Mail: katharina.stowasser-bloch@meduniwien.ac.at

Reihe Qualität - Evaluation - Akkreditierung. Praxishinweise zu Verfahren und Methoden

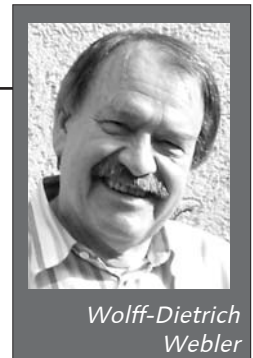
im Verlagsprogramm erhältlich:

Stefanie Schwarz, Don F. Westerheijden, Meike Rehbürg (Hg.): Akkreditierung im Hochschulraum Europa

Bielefeld 2005, ISBN 3-937026-36-3, 261 Seiten, 34,00 Euro

Bestellung - E-Mail: info@universitaetsverlagwebler.de, Fax: 0521/923 610-22

Wolff-Dietrich Webler



Formative und summative Evaluation schärfer unterscheiden! Anmerkungen über den Unterschied zwischen Evaluation und Rückmeldung

Anlass des Artikels

In regelmäßigen Abständen werden anlässlich der Bewertung von Lehrveranstaltungen durch Studierende drei Themen aufgeworfen:¹ 1) Die Frage studentischer Urteilsfähigkeit für derartige Veranstaltungen; 2) Die Kritik an der mangelnden Objektivität solcher Urteile; damit verbunden ist zumindest eine Unklarheit in der Unterscheidung von Evaluation und Rückmeldung (feedback) sowie von summativer und formativer Evaluation; und 3) die verbreitete Focussierung auf die Qualität der Lehre als Indikator für die Qualität des ganzen Studiums (u.U. ergänzt um Studiendauer und Studienabbrecher und -erfolgsquoten).

1. Angeblich mangelnde Urteilsfähigkeit der Studierenden

Lehrende stellen immer wieder fest, dass ihre Intentionen für den bevorstehenden Lehr-/Lernprozess von Studierenden nicht verstanden oder nicht geteilt werden. Wenn die studentische Veranstaltungsbewertung dann nicht so positiv ausfällt wie erhofft, werden Fragen bezüglich der Urteilsfähigkeit der Studierenden aufgeworfen. Andere Fragen – wie z.B. inwieweit die Veranstaltung den individuellen Lernbedürfnissen entgegenkam (m.a.W. wie viel die Lehrenden über studentisches Lernen und Lernbedürfnisse wissen und dies didaktisch berücksichtigen) oder warum Studierende gelangweilt auf den Verlauf einer Veranstaltung reagieren und durch andere Beschäftigungen stören, schließlich, wie weit es Lehrenden gelungen ist, aus der eigenen „Lehr-“Veranstaltung der Lehrenden eine gemeinsame „Lern-“Veranstaltung von Lehrenden und Studierenden zu machen – bleiben oft ausgeblendet.

Studierende bewerten deren Qualität, seit es Lehrveranstaltungen gibt. Dies geschieht meist im Gespräch mit anderen Studierenden, informell, nach subjektiven Maßstäben des Gefallens, der Zufriedenheit. In dieses Gefallen bzw. die Zufriedenheit gehen Urteile darüber mit ein, ob sie den Eindruck haben, neugierig geworden zu sein sowie in dieser Veranstaltung leicht(er) bzw. besser zu lernen (als z.B. in anderen). Und genau das macht ihre Expertise aus: Zu beurteilen, ob die Veranstaltung für sie persönlich, mit ihren individuellen Lernbedürfnissen und Lerngewohnheiten lernförderlich ist oder nicht. Da Lehrerfolg sich (den Zielen einer Veranstaltung gemäß) letztlich ausschließlich am Lernerfolg der Studierenden

bemisst, sind dies Informationen, die außer den Betroffenen sonst niemand liefern kann. An der vorstehenden Tatsache kommt niemand vorbei. In der allgemeinen Hochschul- und Fachliteratur ist lange über die Urteilsfähigkeit von Studierenden gestritten worden. Dabei hält sich hartnäckig die These, dass die Studierenden zu einer Beurteilung der Veranstaltung (allein) nicht fähig seien, ihr Urteil also mindestens relativiert werden muss.

2. Kritik an der mangelnden Objektivität studentischer Urteile – Zur Forderung nach „Objektivität“ der Ergebnisse

2.1 Löst eine „objektive“ Beurteilung von Veranstaltungen einen evtl. eingeschränkten Lernerfolg?

Häufig weisen Lehrende auch auf die Subjektivität der Bewertungen durch die teilnehmenden Studierenden hin. An den meisten Hochschulen wird sich aber entweder mit dem Grad der Objektivierung durch die oft große Zahl von Studierenden und ihre jeweiligen Perspektiven zufrieden gegeben und zugestanden, dass eine Rest-Subjektivität unauflösbar hinzunehmen ist (s.u.). An weiteren Hochschulen wird die Subjektivität der Bewertungen durch die teilnehmenden Studierenden nicht nur für unvermeidlich gehalten, sondern auch für sinnvoll angesichts der Subjektivität der individuellen Lernbedürfnisse und Lernstile (s.u.). In Wien also wird an der dortigen Hochschule u.a. versucht, eine Objektivierung der Bewertung zu erreichen.

Manche Fachautoren stufen (wie in dem Wiener Projekt beschrieben) die „Bewertung der Lehrqualität durch die studentische Einschätzung“ als zu einseitig ein, „da Studierende z.B. die Interaktionsebene zwischen Lehrenden und Studierenden nicht umfassend in ihrer Gesamtheit bewerten können“, wobei sich z.B. auf Reischmann 2006, S. 144, Bachmann 2009, S. 7 sowie Flick 2008, S. 25 bezogen wird. Diese Kritik der Autoren versäumt allerdings, summative von formativer Evaluation zu unterscheiden. Sie klingt nach Forschungsinteresse und allenfalls summativer Evaluation und übersieht den Feedback-Charakter von üblichen Veranstaltungsbewertungen, die der Subjektivität nicht enttrinnen können und

¹ Nun wieder durch den Beitrag von Isolde Kirnbauer und Katharina Stowasser-Bloch über ihr Pilotprojekt an der Medizinischen Universität Wien zur Veranstaltungsbewertung in ds. Heft.

wollen. Wie erwähnt, wird dort angeführt, a) die Studierenden seien vor allem mit dem zu Erlernenden beschäftigt; b) sie könnten im universitären Alltag akzeptierte Routinen nicht mehr wahrnehmen und c) sie würden über ein zu geringes didaktisches Grundwissen verfügen. Solche Gesichtspunkte sind nur dann relevant, wenn eine „objektive“ Beurteilung gesucht wird. Dann gilt aber gleichlautend: a) viele Lehrende sind überwiegend mit den zu lehrenden Inhalten beschäftigt; b) die Lehrenden können im universitären Alltag akzeptierte Routinen nicht mehr wahrnehmen und c) verfügen sie im Regelfall – solange „Lehren lernen“ nicht zur professionellen Ausbildung der Lehrenden an Hochschulen gehört – über ein zu geringes didaktisches Grundwissen.

Insbesondere wenn Lehrende den Ausgang studentischer Veranstaltungsbewertung nicht nachvollziehen können, kommt der Wunsch nach einer „objektiveren“ Beurteilung der Lehrleistung auf, eine, die nicht den Lernerfolg als entscheidendes Kriterium – selbstverständlich bei adäquaten Lehrzielen – zur Grundlage nimmt.

Dahinter steht die Weigerung, Beurteilungsmaßstäbe der Studierenden anzuerkennen und die Absicht, sie (zumindest teilweise) durch eigene zu ersetzen. Dabei geht es nicht um „Gefallen“ oder „Zufriedenheit“ mit einer Veranstaltung. Das ganze zu einem „Objektivierungsproblem“ – also einem Methodenproblem – zu machen, geht an dem Kern des Konflikts vorbei. Dort geht es nicht um Studierende und ihre Erfolgskriterien, sondern um die Maßstäbe bei der Beurteilung von Lehr- und Lernveranstaltungen, also um den Qualitätsbegriff. Ein solcher Konflikt müsste dann mit den Konstrukteuren des Veranstaltungsfragebogens ausgetragen werden. Denn jeder etwas komplexere Fragebogen spiegelt unvermeidlich die Qualitätsvorstellungen der Urheber wider. Dort wird nach Merkmalen guter Lehre als Bestandteile guten Studiums gefragt und nach Prozessvariablen, mit denen die Interaktionsprozesse künftig optimierend beeinflussbar sind. Die Studierenden haben nur in den freien Kommentaren die Möglichkeit, eigene, im Fragebogen möglicherweise vernachlässigte Aspekte einzubringen. Und eine solche Auseinandersetzung mit den Fragebogenkonstrukteuren um die Qualitätskriterien guter Lehr- und Lernveranstaltungen lohnt.

Die Unzufriedenheit von Lehrenden liegt oft an der mangelnden Qualität der Instrumente, mit denen den Lehrenden eine Rückmeldung über die Lernförderlichkeit ihrer Veranstaltung verschafft werden soll (vgl. Ziff. 4). Die Art wie und was über eine Veranstaltung erfragt wird, hat oft zu Wünschen übrig gelassen. Zum Beispiel hat die häufig gestellte Frage „Der Dozent ist fachlich kompetent“ in solchen Fragebögen nichts zu suchen, da sich die Studierenden sicherlich nicht ausreichend mit dem wissenschaftlichen Werk der betreffenden Person auseinandergesetzt haben (von einer solchen fachlichen Urteilsfähigkeit ganz abgesehen).

Der Versuch, ein objektives Maß an Güte der Veranstaltung zu finden („Lehrqualität“), das sich nicht am Lernerfolg der Anwesenden orientiert, ist also kritisch zu sehen. Lernerfolg ist nun einmal in seinem Zustandekommen subjektiv geprägt. Die Vorstellung, es könnte eine objektiv gute Veranstaltung geben, auch wenn die

Veranstaltung von Studierenden aus ihrer subjektiven Sicht abgelehnt wird, ist unzutreffend.²

„Objektiv – vom medizinischen Standpunkt aus betrachtet – wurde alles richtig gemacht; die Operation war erfolgreich. Aber der Patient ist tot.“ Es gibt keine objektive Beurteilung, die die Beurteilung der Studierenden aushebeln könnte, weil die Studierenden aufgrund ihrer Defizite die Veranstaltung nicht richtig verstanden hätten und deshalb falsch bewertet hätten.

Dass Studierende (vor allem der Anfangssemester) vielfach keine oder falsche Lernstrategien haben und dies durch zusätzliche Angebote in- oder außerhalb der Fachveranstaltungen bearbeitet werden muss, wird dadurch nicht ausgeschlossen. Auch hier gilt, dass die Hochschule sich ihre Studierenden nicht ideal modellieren kann, sondern Lernerfolge mit den real existierenden Studierenden und mit deren jeweiliger Lerngeschichte erzielen muss.

Nur wenn Lehrerfolg von Lernerfolg trennbar wäre, könnte es eine „objektive Lehrqualität“ geben (die vom Lernerfolg absieht und sich an anderen Kriterien orientiert). Dann wären wir wieder bei dem Punkt: „Die Studierenden haben selbst Schuld, wenn sie bei mir nichts lernen!“ Oder (tatsächlich Originalton eines Hochschullehrers): „Wer meine Vorlesung nicht versteht, gehört nicht an eine Universität!“ Lehrerfolg kann nur von „Gefallen“ und (erster) Zufriedenheit getrennt werden, aber nicht vom Lernerfolg.

Objektivität ist das eine – wenn das Lehrangebot aber bei den Lernern nicht ankommt, hat es seinen Zweck verfehlt, auch wenn Dritte die Lehre in der Veranstaltung für überzeugend erklären..

2.2 Formative oder summative Evaluation?

Veranstaltungsbewertungen werden oft am Semesterende eingesetzt. Dann ist das Semester vorüber, eine Chance zur Verbesserung dieser konkreten Veranstaltung besteht nur für den Wiederholungsfall. Also hat die Befragung die Funktion, das Semester zu bilanzieren. Wenn auch noch die bekannt gegebenen Ziele (in einem kommentierten Veranstaltungsverzeichnis bzw. dem Modulhandbuch) explizit oder implizit dazu genommen werden, handelt es sich um eine summative Evaluation – wenn, und nur dann, wenn auch der Grad der Zielerreichung erhoben wurde. Steht eher eine laufende Optimierung der Lehr-/Lernprozesse im Mittelpunkt, dann ist es sinnvoller, die Befragung schon zur Mitte des Semesters, also noch im Prozessverlauf, durchzuführen (schnelle Auswertung voraus gesetzt). Dann besteht die Chance, aufgrund der Befragungsergebnisse noch Änderungen vorzunehmen (Feedback-Funktion der Befragung, evtl. formative Evaluation).

Je nach Verwendungszweck der Daten stellen sich z.T. unterschiedliche Fragen zum Befragungsinstrument und zum Verfahren. Daher muss früh entschieden werden, für welche Zwecke die Befragungen eingesetzt werden

² Solche Ausnahmen betreffen negative studentische Urteile, weil Studierende sich mehr Konsum statt persönlicher Anstrengung („Lernarbeit“) gewünscht hätten oder dass manche Lehrveranstaltungen konfliktuell verlaufen und dadurch intensivere Beschäftigung mit den Inhalten auslösen („wer hat denn nun eigentlich recht“) als eine harmonische Spaßveranstaltung.

sollen. Beide Evaluationsformen werden aber oft verwechselt. Im einen Fall – der summativen Evaluation – haben die studentischen Veranstaltungsbewertungen – bei relativ feststehenden Zuordnungen von Veranstaltungen zu Lehrenden über einen längeren Zeitraum – eine Studienberatungsfunktion, werden also – vor allem in den USA – als Hilfe bei der Wahl von Veranstaltungen verwendet. Sie werden vor allem aber auch zur Personalbeurteilung eingesetzt, zur Einstufung der Lehrkompetenz der Lehrenden mit Konsequenzen für ihre Laufbahn (Vertragsverlängerung, Übernahme in ein unbefristetes Vertragsverhältnis – tenure – oder Gehaltshöhe usw.). Dies geschieht in den USA, Canada und einzelnen europäischen Universitäten. Dabei sind in den USA und Canada die abgegebenen Studierendenbewertungen nicht die alleinige Grundlage. Die dort üblichen Formen des „student rating“ waren Auslöser für die Entstehung der „teaching portfolios“ der Lehrenden, einer Kombination aus Tätigkeitsbericht und Darstellung der eigenen „teaching philosophy“, die dem student rating die Perspektive der jeweiligen Lehrenden gegenüber stellen sollte – aber nur als Reflektionshilfe und bei Personalbeurteilungen.

Im anderen Fall werden die studentischen Veranstaltungsbewertungen als Rückmeldeschleifen zur Optimierung des laufenden Lehr-/Lernprozesses mit den konkret Beteiligten verwendet. Mit dieser Funktion werden sie überwiegend an europäischen Hochschulen eingesetzt. Die studentischen Bewertungen müssen bzgl. ihrer Berechtigung im Licht vor allem der Studienziele, optimierter Lernbedingungen und der Studienordnung interpretiert werden (was in einem mündlichen Dialog zwischen Lehrenden und Lernenden einer Veranstaltung üblicherweise auch geschieht). Wird dies z.B. im Rahmen einer systematischen Begleitforschung (etwa für einen neuen Studiengang oder für die Neuerprobung eines Teilcurriculums) mit wiederholten Erhebungszeitpunkten eingesetzt, dann handelt es sich um einen anderen Evaluationstyp, die formative Evaluation.

Während also die formative, prozessbegleitende Evaluation auf eine Verbesserung der Lehr-/Lernprozesse durch laufende bzw. periodische Rückkoppelung von Prozessdaten gerichtet ist, beruht die summative Evaluation auf einer punktuellen Bilanz und hat das Ziel, ein abschließendes Urteil zu fällen. Letzteres wäre als Abschluss eines Habilitationsverfahrens zur Erteilung der *venia legendi* denkbar oder auch bei einer Probeveranstaltung in einem Berufungsverfahren. Aber im Alltag der Lehre wird die prozessoptimierende Variante vorgezogen.

Aus der Sicht traditioneller empirischer Sozialforschung wird gegen die formative Evaluation eingewandt, dass der Untersuchungsgegenstand durch die Rückkoppelung der Daten in dem (in diesem Modell ausdrücklich gewollten) permanenten Lernprozess ständig verändert wird. Die Kritik erinnert an die Kritik an der in den 1970er Jahren aufkommenden Aktionsforschung. Die Dateninterpretation am Schluss ergebe keine verlässliche Zustandsbeschreibung, man wisse eigentlich nicht, welche Erkenntnisse über den Gegenstand fixierbar seien. Diese klassische Konfliktsituation kann sich noch einmal zuspitzen, wenn die Bewährung eines Modells durch Begleitforschung festgestellt und je nach Ergebnis

allgemein für verbindlich erklärt werden soll. Das kann z.B. in der Schulentwicklungsforschung der Fall sein, das war in einer besonders zugespitzten Situation bei der bundesweiten, auch zwischen 7 Bundesländern konkurrierenden Erprobung von Modellen einstufiger Juristenausbildung Ende der 1970er/Anfang der 1980er Jahre der Fall. Bei dieser Erprobung bestanden die 7 an dem Modellversuch beteiligten juristischen Fachbereiche darauf, ihre Modelle mit Hilfe der gewonnenen empirischen Daten in laufender Rückkoppelung verbessern zu können (formative Evaluation), während das Bundesjustizministerium eine Evaluation verlangte, die die Modelle 6 Jahre in unverändertem Zustand begleiten und dann über die Ergebnisse der fixierten Modelle berichten sollte (summative Evaluation)(vgl. Webler 1981, 1983a, 1983b).

3. „Qualität der Lehre“ oder „Qualität des Studiums“? Ziele und Zwecke von Lehrveranstaltungen und deren Qualität als Untersuchungsgegenstand

Wenn der Präsident der deutschen Hochschulrektorenkonferenz anlässlich der Erinnerung an die vor 10 Jahren begonnenen Bologna-Reformen betont, Studium müsse Persönlichkeiten hervorbringen, nicht nur Absolventen, und das gelinge in dem verkürzten Bachelor-Studium nicht; wenn gleichzeitig Steigerungen der Studienerfolgsquote infolge der Bologna-Reform öffentlich als Erfolge gefeiert werden, dann liegen hier offensichtlich fundamental unterschiedliche Qualitätsvorstellungen zugrunde, die sicherlich bis in die Bewertung von Lehrveranstaltungen als Basis hinein reichen. Auf eine kurze, plakative Formel gebracht: Geht es um das Drillen von Fachexperten oder die Bildung von Menschen als Ziele des Vorhabens?

Qualität des Studiums

Die Studierenden kommen an die Hochschule nicht in erster Linie, um Lehrveranstaltungen zu besuchen, sondern um ein Studium zu absolvieren. Das wird häufig (nahezu) gleichgesetzt, ist aber etwas deutlich anderes. Studium stellt einen kompakten Sozialisationszusammenhang dar, an dessen Ende sich Akademiker/Akademikerinnen entwickelt haben. Das können die Studierenden allein in Lehrveranstaltungen nicht werden. Studium besteht aus einer Fülle formeller und informeller Lernorte und -situationen, von denen Lehrveranstaltungen nur einen relativ kleinen Teil ausmachen. Soll die Güte einer Hochschule als Studienort, die Güte eines Studiengangs für ein gutes Studienergebnis festgestellt werden, gehört weit mehr dazu, als die Lehrveranstaltungen zu evaluieren.

Falls also in einem QM-Programm von den Maßnahmen her nur die „Qualität der Lehre“ festgestellt, gesichert und gesteigert werden soll, dann muss zunächst der Stellenwert der Lehre im Gesamtverbund *des Studiums der Studierenden* geklärt werden, weil die sofortige Betrachtung der Lehre und ihrer Qualität (und dann nur noch von förmlichen Lehrveranstaltungen) leicht zu einer Perspektivenverengung der Betrachtung des Studi-

ums führt.³ Eine Verselbständigung der Qualität von Lehrveranstaltungen könnte sich leicht als kontraproduktiv im Sinne der Qualitätssteigerung des Gesamtstudiums erweisen.

Was bedeutet „Studium“? An Hochschulen bedeutet es die Auseinandersetzung mit Wissenschaft (meist anhand einer Wissenschaftsdisziplin) in Selbststudium und Lehrveranstaltungen und in studentischen Subkulturen (Freizeit, Semesterfeiern, Wohnheimen, studentischem Breitensport, studentischer Selbstverwaltung usw.). In diesem umfassenden Sozialisationsprozess wird sich aktiv formell und vor allem informell gebildet. Das sich immer wieder klar zu machen ist auch im Rahmen der Debatte um Kompetenzerwerb immer wichtiger geworden, weil der Entwicklung vieler Studiengänge kein Gesamtbild der gewünschten Absolvent/innen und ihrer Einstellungen, Fähigkeiten und Fertigkeiten zugrunde liegt. Die Überlegung muss von diesem Gesamtbild (und ihren individuellen Variationen) ausgehend überprüfen, wo – an welchen Lernorten im Gesamtstudium – die gewünschten Einstellungen, Fähigkeiten und Fertigkeiten erworben werden können. Zielvorstellungen, Profile und die von der Hochschule direkt beeinflussbaren Lernorte müssen soweit möglich aufeinander abgestimmt werden. Eine Anstalt zum Drill von Fachexperten wird niemals die Effekte erzielen, die von einer „akademischen Bildung“ erwartet werden. Die Hochschulen sollen also *das Studium fördern*, d.h. *Lernprozesse der Studierenden ermöglichen*, z.T. auch organisieren. Es geht also immer um die Qualität des Studiums und nur als deren Teilmenge um die Qualität der Lehre.

Qualität der Lehre

Wird nach der Qualität der Lehre gefragt wie in dem Wiener Beispiel, muss als erstes der Qualitätsbegriff geklärt werden. Der ist zwar auch als DIN ISO-Norm festgelegt, aber eindrücklich knapp lautet immer noch die britische Formulierung: „Quality means fitness for purpose“ (Ball 1985). Also muss zur Beurteilung von Lehrveranstaltungen deren „purpose“ festgestellt werden. Da geht es um die Gesamtziele des Studiums, zu denen jede Lehrveranstaltung Anteile beiträgt, erkennbar am outcome, also nicht dem, was behandelt wurde (input), sondern dem, was die Studierenden anschließend können (sollen) (outcome). **Es muss klar sein, welche Ziele und welchen Zweck Lehrveranstaltungen verfolgen** sollen, um ihre Qualität feststellen zu können. Ein Qualitätskriterium besteht darin, inwieweit die Lehrveranstaltungen die Studierenden zum Selbststudium befähigen. Im Kontext eines vollwertigen Studiums ist dies unverzichtbar, aber ebenso als Vorbedingung für lebenslanges Lernen.

Die Hauptkonfliktpunkte liegen in der ständigen Verkürzung der Lehrziele auf Wissensziele – ein gravierender Qualitätsmangel – und in den mangelnden Kenntnissen rund um das Lernen der Studierenden. Deshalb bleiben viele Veranstaltungen bei der Frage einer optimierten Präsentation ihres Stoffes stehen, weil sie zur Frage der besten Art der Aneignung und Verarbeitung durch die Studierenden nicht vordringen können. Und genau dort sollte ein Schwerpunkt der Fragen in den Veranstaltungsfragebögen liegen.

Ziel guter Lehre muss es sein, immer unter Orientierung an den übergeordneten Lehrzielen die Veranstaltungen enger an den Lernbedürfnissen und -möglichkeiten der Studierenden auszurichten. Da gibt es bereits zahlreiche Beratungsformate entweder mit Hilfe einer Hospitation durch Personen mit praktischer hochschuldidaktischer Expertise (also auch hochschuldidaktisch geschulte Lehrende) in der Veranstaltung oder (gute Erhebungsinstrumente voraus gesetzt) in der nachträglichen Interpretation der studentischen Erhebungsdaten.

4. Qualität des Befragungsinstruments

Was können Befragungen von Studierenden leisten? Um es zu wiederholen: Sie können hauptsächlich feststellen, in welchem Maße Grundanlage und realer Verlauf einer Veranstaltung den Lernbedürfnissen der Studierenden unterstützend entgegen kamen. Das ist die Kernaussage, und das ist eine Information, die ohnehin nur die Studierenden liefern können. Niemand kann sie dort ersetzen. Eine förmliche „studentische Lehrendenbewertung“, wie in dem Wiener Pilotprojekt, sollte es nicht geben. Es geht nicht darum, Lehrende zu bewerten (und schon gar nicht personifiziert) sondern eine gemeinsame Veranstaltung, und zwar eine LERNveranstaltung. Eine solche Begrifflichkeit und ein solcher Zuschnitt (nicht die gemeinsame Interaktion wird beurteilt, sondern die Lehrperson) suggerieren, dass die Lehrenden für das Gelingen einer Veranstaltung allein die Verantwortung tragen. Das genau ist nicht der Fall.

Der Wunsch, den Lernerfolg am Ende einer Semesterveranstaltung nicht nur näherungsweise festzustellen, sondern sogar präzise zu messen, ist sehr verbreitet (vgl. Reischmann 2006). Aber dieser Wunsch verkennt die Langfristigkeit komplexerer Lernprozesse, auf die es im Studium ankommt und die Tatsache, dass Lernprozesse häufig längere Zeit verdeckt, d.h. zunächst unbewusst verlaufen. Dadurch können sie nicht abgefragt werden. Ein Ereignis, ein Zusammenhang, ein Streit beschäftigen die Person unbewusst weiter. Mosaikartig kommen weitere Eindrücke (z.B. aus weiteren Lehrveranstaltungen und aus informellem Lernen) dazu. Und plötzlich erschließt sich ein Ergebnis, wird ein komplexer Zusammenhang begriffen („mir geht ein Licht auf“ oder ein ganzer Lichtenbaum). Wenn dies in einer Lehrveranstaltung geschieht, wird diese letzte Veranstaltung in Befragungen überbewertet, dazwischen liegende, unverzichtbare, aber verdeckte Stationen des Lernens u.U. unterschätzt. Befragungen der Studierenden (oder unmittelbar anschließende Klausuren) können daher allenfalls zu einem Teil die Lernerfolge wiedergeben, die (zumindest in ihrer vollen Dimensionalität) erst langfristig und als Gemeinschaftsleistung vieler Veranstaltungen und anderer Lernorte eintreten. Daher kommt auch ein Teil der Kritik an studienbegleitenden Prüfungen – noch dazu solchen, deren Ergebnisse bereits in die Endnote eingehen. Lernerfolge einiger Komplexität zu erfassen, ist eher Gegenstand von Absolventenstudien.

³ Deswegen hat der Verfasser auch einen provokanten Artikel veröffentlicht mit dem Titel „Gebt den Studierenden ihr Studium zurück!“ (Webler 2005)

Es war schon davon die Rede, dass in jeden Fragebogen, mit dem Lehrveranstaltungen bewertet werden sollen, implizit ein didaktisches Modell guter Lehre – nämlich das der Urheber – eingeht. Insofern beginnt die Subjektivität des Unternehmens schon bei der Fragebogenkonstruktion. Beim Vergleich verschiedener Veranstaltungsfragebögen ist immer wieder zu beobachten, dass dort relativ schlichte didaktische Modelle für den Aufbau und dann entsprechend für die Wahrnehmung einer Veranstaltung eingearbeitet werden. Das begrenzt die Verwendbarkeit der Ergebnisse für eine eigene Weiterbildung der betreffenden Lehrenden erheblich. Ein Fragebogen, der wie häufig anzutreffen, die Bewertung der Lehrenden in den Mittelpunkt stellt, sendet implizit die Botschaft an die Studierenden aus; „Für meinen Lernerfolg, für die Qualität der Veranstaltung insgesamt ist allein die lehrende Person verantwortlich“. Daher ist es immer wieder bei der Analyse derartiger Fragebögen aufschlussreich, wie viele Fragen die lehrende Person und ihr Verhalten betreffen im Verhältnis zu den anderen Variablen einer Lehrveranstaltung. Ein anspruchsvoller Fragebogen sollte als Dimensionen a) einige Rahmenbedingungen, b) die Transparenz der Ziele (gewünschten Lernerfolge in Kompetenzen), c) die Lehrendenseite, d) die Studierendenseite und dann vor allem e) die Interaktion zwischen beiden (und die dafür eingesetzten Methoden) enthalten. Schon eine solche Dimensionierung (gestützt durch die Art der Fragestellungen) macht deutlich, dass der Erfolg einer Veranstaltung von beiden Seiten (und allerdings auch den vorausgegangenen didaktischen Entscheidungen) abhängt. Entsprechend differenziert werden die Aussagen der Studierenden sein.

In dem Wiener Beispiel wird die Interaktion der Studierenden, werden evtl. „Störungen“ durch die externen Beobachter registriert. Das Alles gehört als Wahrnehmungsdimensionen in einen studentischen Fragebogen hinein. Im Bielefelder Fragebogen (1991/92 entwickelt und als Fortschreibungen in viele lokale Fragebogenentwicklungen mit eingegangen), haben die Studierenden einer Lehrveranstaltung selbstverständlich (neben Fragen zum eigenen Studienverhalten) auch das Studienverhalten der Kommilitonen beurteilt und dabei auch bewertet, ob sie dadurch angeregt, gestört usw. wurden. Als Beispiel einer item-Batterie daraus:

18.	Mitstudierende
	<ul style="list-style-type: none"> • Gespräche unter den Studierenden stören die Veranstaltung • In der Veranstaltung herrscht häufig ein Kommen ein • Das Klima unter den Studierenden ist kooperativ • Die Studierenden wirken gut vorbereitet • Die Studierenden beteiligen sich angeregt • Die Studierenden zeigen Interesse

Hier stellt sich zunächst die Frage, wofür die abgefragten Merkmale stehen. In der Bewertung gilt es dann zu entscheiden, ob die Einschätzungen als lernförderlich oder als hinderlich einzustufen sind.

Die Beobachterinnen registrieren „Störungen des Unterrichts“. Abgesehen davon, dass es sich im Hochschulstudium nicht um „Unterricht“ im Alltagssprachlichen Sinn handeln sollte, ist auch der Begriff „Störung“ zu hinterfragen. Aus der Schulunterrichtsforschung ist bekannt,

dass viele Schülergespräche, die von den Lehrer/innen als Disziplinproblem wahrgenommen werden, sich in Wahrheit um Unterrichtsthemen drehen, also entweder der Verarbeitung des Unterrichtsstoffes dienen oder Unklarheiten im Verständnis betreffen (die natürlich dann im Plenum gefragt werden sollten, um sie klären zu können. Aber dazu fehlt dann oft der Mut). Sie sind eher als produktiv für das Lernen zu interpretieren. In großen Hörsälen allerdings ist der Geräuschpegel oft auch ein Zeichen für die Anonymität der Lernsituation, also ein Defizitmerkmal der Rahmenbedingungen.⁴

5. (Eigen-)Bewertung der Veranstaltung durch die Lehrenden

Dabei kann es – aber aus Gründen der Weiterbildung – sinnvoll sein, die Lehrenden – wie in dem Bielefelder Modell – selbst einen ähnlichen Fragebogen über die Veranstaltung ausfüllen zu lassen wie die Studierenden, um anschließend mit ihnen Selbstbild und Fremdbild bzgl. der Wahrnehmungsunterschiede besprechen zu können. Dabei geht es immer wieder darum, die eigenen Lehrintentionen mit dem zu vergleichen, was und wie es bei den Studierenden ankommt.

Zunächst für Lehrende: Es mag sein, dass dieser Fragebogen zeitgleich mit der studentischen Veranstaltungsbewertung eingesetzt wird, meist also gegen Ende des Semesters oder sogar an dessen Ende. Er sollte professionell gestaltet sein (so jedenfalls ist es zu wünschen) und mit dem Bogen der Studierenden zugunsten der Vergleichbarkeit der Eindrücke als Fremd- und Selbstbild in weiten Teilen identisch sein. Indem die Lehrenden ihre bisherigen Eindrücke und Beobachtungen im Semesterverlauf noch einmal reflektieren, denken sie auch über die Differenz zwischen ihren eigenen Erwartungen und dem tatsächlichen Semesterverlauf (jedenfalls ihrem Bild davon) nach. Sie denken damit auch über ihre Ziele und deren Angemessenheit, über deren Berechtigung nach. Mit ihren Erwartungen setzen sie Maßstäbe. Je weiter diese Erwartungen und das real beobachtete Studierendenverhalten auseinander liegen, desto mehr werden sie über mögliche Gründe für diese Differenz nachdenken. Erfahrungsgemäß reicht das von anfänglich sehr einfachen Schuldzuweisungen bis zu ansteigend selbstkritischen Überlegungen.

Indem sie aufgefordert sind (gestützt durch die Fragen im Fragebogen, die sie auf unterschiedliche Aspekte noch einmal aufmerksam machen) über ihre Erwartungen an Studienverhalten nachzudenken, kommen sie (in ihrer Weiterentwicklung als Lehrpersönlichkeit) auch an den Punkt, an dem sie sich fragen, ob sie eigentlich lernunterstützend lehren. Zum Teil kommen sie angesichts des realen Studienverhaltens und dessen mehr oder weniger zielführender Art auch zu dem Schluss, hier müsste – neben Änderungen am eigenen Lehrkonzept – auch z.T. eine Lernberatung einsetzen, zu der sie mangels Ausbildung ebenfalls nicht in der Lage sind. Dieses ist

⁴ Das Alles ist in viele lokale Fragebogenentwicklungen mit eingegangen – allein schon über die dann an anderen Hochschulen das Qualitätsmanagement aufbauenden Mitarbeiter des damaligen Projekts, Volker Domeyer, Bernd Schiebel, Werner Link und Michael Craanen.

auch der ideale Zeitpunkt für Beratungen, weil viele Lehrende mit didaktischen Schlussfolgerungen aus den (dann oft auch noch widersprüchlichen) Antworten der Studierenden überfordert sind.

Außerordentlich wirksam für einen Lernprozess ist es regelmäßig, zu irritieren, also auch das Studienverhalten der Studierenden durch die Lehrenden bewerten zu lassen (Webler 2005). Die Wirksamkeit gilt für beide – Lehrende wie Lernende.

Dann auch Rückmeldung an die Studierenden: Die Studierenden haben sich zunächst ganz an die Perspektive gewöhnt, „die Lehrenden sind allein für das Gelingen von Veranstaltungen verantwortlich“.

Das ist noch dazu bei unprofessionell gestellten Fragen im Bewertungsbogen für die Studierenden verstärkt worden, die eine solche einseitige Perspektive befördern (Webler 2008). Nun werden sie durch die Bewertung ihres Studienverhaltens durch die Lehrenden daran erinnert, dass Lehrveranstaltungen ein wechselseitiger kommunikativer Prozess sind, an deren Gelingen sie wesentlichen Anteil haben (wenn sie nicht durch Lehrverhalten zu stummen Zuhörern und Zuschauern der „Performance“ der Lehrenden werden). Das ist für sie Anlass, anhand der geäußerten Erwartungen und Eindrücke der Lehrenden selbst über ihre Situation im Studium nachzudenken und Rollenklärung zu betreiben. Zwar liegt zwischen diesen Einsichten und deren praktischer Realisierung durch geändertes Studienverhalten oft noch ein langer Weg, aber ein Anfang ist gemacht. Die Besprechung kann vielleicht synoptisch bei Präsentation ihrer eigenen ausgewerteten Befragungsergebnisse in der Veranstaltung geschehen.

Viele Lehrende sind dazu übergegangen, diesen Fragebogen zu Beginn des Semesters als Instrument zur strukturierten Bekanntgabe eigener Erwartungen und Lehrziele (wichtig für die Studierenden) und am Ende des Semesters (als Feedback) erneut einzusetzen. Die Studierenden haben dies nach anfänglichem Erschrecken, selbst beurteilt zu werden, als sehr nützlich aufgenommen.

6. Triangulation der Beobachtung? Triangulation der Bewertung? Oder keines von beiden?

Nun existiert die Forderung, Sachverhalte in Lehrveranstaltungen durch Triangulation zu beobachten, um zu objektiven Urteilen zu kommen. Der Begriff entstammt der Mathematik und bedeutet in der Sozialforschung die Betrachtung eines Vorgangs, eines (Interaktions-)Prozesses durch drei verschiedene Akteure aus unterschiedlichen Richtungen. Es geht also zunächst um ein Mittel der Objektivierung der Wahrnehmung bei der Dokumentation komplexer Sachverhalte. Erst im zweiten Schritt geht es um die Bewertung. (Im Fußball wird dieses Problem gerade bei der Feststellung eines erzielten Tores diskutiert; im Hockey schon der Video-Beweis angewandt). Die anschließenden Regeln für die Interpretation des Ergebnisses sind im Sport klar und werden von den Beteiligten geteilt. Das Prinzip liegt auch dem in der Wissenschaft sehr vertrauten Vorgang der Begutachtung zugrunde. Zwei Gutachter werden eingesetzt, um zu si-

chern, dass das beschriebene Forschungsvorhaben oder eine Dissertation/Habilitationsschrift in ihrer Komplexität möglichst zuverlässig erfasst wird. Auch hier wird idealiter bei der anschließenden Beurteilung von der Anwendung übereinstimmender wissenschaftlicher Standards ausgegangen. Erst bei Uneinigkeit wird ein dritter Gutachter eingesetzt. Die Video-Dokumentation wird bei Lehrveranstaltungen längst eingesetzt, aber als unmittelbares Feedback-Instrument für Weiterbildungszwecke.

Die Triangulation ist für den Dokumentationszweck sehr aufwändig, weil drei Personen den Vorgang in seiner vollen Länge beobachten und protokollieren müssen. Dem muss eine Schulung für die Beobachtung vorausgegangen sein. Als weiterer, aufwändiger Vorgang kommt dann die Abgleichung der Protokolle hinsichtlich der dokumentierten Prozesse hinzu. Die Video-Dokumentation ist für den Dokumentationszweck dann bei weitem vorzuziehen. Aber die Triangulation in dem Evaluationsprojekt bezieht sich vor allem auf die Bewertung der Vorgänge, also letztlich auf die Frage nach der Qualität der Lehr- und Lernveranstaltung. Die Dokumentation könnte viel leichter über Videoaufnahmen geleistet werden, die Frage von Bewertungskonflikten ist nicht durch Prozessbeobachtung, sondern durch Klärung vorab zu lösen.

7. Aufwand, Kosten, Angemessenheit

Der organisatorische Aufwand, die Vorbereitung (Schulung bzw. Weiterbildung der Beobachter/innen) und die Kosten einer Triangulation sind erheblich. Eine Qualitätssteigerung wird nur dann erreicht, wenn die Maßstäbe professionell sind und die Beteiligten entsprechend anspruchsvoll ausgebildet sind.

Bei Einsatz von Beobachter/innen müssen diese in der Beobachtung und Bewertung hochschuldidaktischer Prozesse geschult sein. Das setzt wieder eine Klärung hochschuldidaktischer Ziele und deren Bewertung voraus, weil diese die Beobachtungssensibilität stark beeinflussen.

Auf die Gesamtheit der Lehrveranstaltungen einer Hochschule (die Universität Bielefeld hatte zu einem bestimmten Zeitpunkt in einem Semester über 800 Lehrveranstaltungen) hochgerechnet, müsste der Erkenntniszuwachs dieser Methode (insbesondere um der Objektivierungsforderung nachzukommen zwei bis drei Beobachter/innen einzusetzen) schon sehr erheblich sein, um die Kosten zu rechtfertigen. Mit der Frage nach der Angemessenheit der Methode stellt sich die Frage, ob es funktional äquivalente Alternativen gibt. Im Regelfall der Veranstaltungsbewertung geht es um eine intern gewünschte Optimierung der Veranstaltung (über Rückmeldung) auf der Basis der Feststellung des status quo, nicht um ein außenwirksames Urteil als summative Evaluation. Von daher überzieht das Verfahren den Anlass und verwechselt die beiden Evaluationstypen. Da eine summative Evaluation der Ausnahmefall sein wird, erscheint der Aufwand zu hoch. Aber vor allem erscheint das Streben nach Objektivität nicht erfolversprechend – es sei denn, die Studierenden würden sich gegen einen Lernprozess sträuben, der im Ergebnis hohe Lernerfolge brächte.

Literaturverzeichnis

- Bachmann; H. W. (2009): Systematische Lehrveranstaltungsbeobachtungen an einer Hochschule. Bielefeld.
- Ball, C. (1985): Fitness for Purpose: Essays in Higher Education. Edited by Dorna Urwin. Guildford and NFER-Nelson.
- Bortz, J./Döring N. (2002): Forschungsmethoden und Evaluation. Berlin, Heidelberg.
- Flick, U. (2008): Triangulation. Eine Einführung. Wiesbaden.
- Reischmann, J. (2006): Weiterbildungsevaluation. Lernerfolge messbar machen. Augsburg.
- Webler, W.-D. (1981): „Begleitforschung und Politische Entscheidung. Das Beispiel der Neuregelung der Juristenausbildung.“ In: Janshen, D./Keck, O./Webler, W.-D. (Hg.): Technischer und Sozialer Wandel. Eine Herausforderung an die Sozialwissenschaften. Königstein/Ts.
- Ders. (1983a): „Politikberatung durch Begleitforschung - Politische und forschungsmethodische Probleme am Beispiel der Juristenausbildung.“ In: Hellstern, G.-M./Wollmann, H. (Hg.): Experimentelle Politik – Reformstrohfeuer oder Lernstrategie? (Beiträge zur sozialwissenschaftlichen Forschung 27), Opladen.
- Ders. (1983b): „Forschungsansätze in der offenen Entwicklung und Evaluation sozialer Experimentalprogramme.“ In: Hellstern, G.-M./Wollmann, H. (Hg.): Experimentelle Politik – Reformstrohfeuer oder Lernstrategie? (Beiträge zur sozialwissenschaftlichen Forschung 27), Opladen.
- Ders. (1991): „Kriterien für gute akademische Lehre.“ In: Das Hochschulwesen, Jg. 39/H.6, S. 243 - 249.
- Ders. (1993): „Evaluation der Lehre: Praxiserfahrungen und Methodenhinweise“. In: Winkler, H. (Hg.): Qualität der Hochschulausbildung. Verlauf und Ergebnisse eines Kolloquiums an der Gesamthochschule Kassel. (Wissenschaftliches Zentrum für Berufs- und Hochschulforschung, Universität Gesamthochschule Kassel, Werkstattberichte 40) Kassel.
- Ders. (1995): „Kriterien für gute akademische Lehre.“ In: Standhart, R/Löhmer, C. (Hg.): Lebendiges Lernen in toten Räumen. Zur Verbesserung der Lehre an der Hochschule. Giessen.

- Ders. (2005): „‘Gebt den Studierenden ihr Studium zurück!’ Über Selbststudium, optimierende Lernstrategien und autonomes Lernen (in Gruppen).“ In: Beiträge zur Lehrerbildung, Jg. 23/H.1, S. 22-34.
- Ders. (2005): „Zur Bewertung von Lernveranstaltungen. Konzeptionelle Begründung des Bielefelder Modells der Evaluation von Lehrveranstaltungen.“ In: Das Hochschulwesen, Jg. 53/H.2, S. 63-70.
- Ders. (2008): „Was taugen Fragebögen zur Lehrveranstaltungsbewertung - und was ihre Ergebnisse? Praxishinweise aus einem Vergleich üblicher Fragebögen zur Evaluation von Lehre und Studium.“ In: Qualität in der Wissenschaft, Jg.3/H.4, S. 96-105.
- Ders. (2009): „Wieviel Wissenschaft braucht die Evaluation? Evaluation von Lehre und Studium als Hypothesenprüfung.“ In: Qualität in der Wissenschaft, Jg.4/H.1/2, S. 28-37.
- Ders. (2011): „Erfassung der Qualität der Lehre – Warum so viele Versuche scheitern.“ In: Hochschulmanagement, Jg.6/H.4, S. 84-93.

■ **Dr. Wolff-Dietrich Webler**, Professor of Higher Education, University of Bergen (Norway); Ehrenprofessor der Staatlichen Pädagogischen Universität Jaroslaw/Wolga; Leiter des Instituts für Wissenschafts- und Bildungsforschung Bielefeld (IWBB), E-Mail: webler@iwbb.de

Studierende an Fachhochschulen: rd. 50% der Zahl der Studierenden an Universitäten erreicht

(QiW) Die quantitativen Größenverhältnisse zwischen den Hochschulen in Deutschland sind seit der Bologna-Strukturreform der Studiengänge gut vergleichbar geworden. Daher lohnt ein aktueller Größenvergleich, der nachfolgend zusammengestellt ist:

Zahl der Hochschulen in Deutschland insgesamt: 333 (Stand 02.08.2012)

Darunter **110 „Universitäten und Hochschulen mit Promotionsrecht“**, davon 88 in staatlicher Trägerschaft, 11 in privater Trägerschaft (staatlich anerkannt) sowie 11 in kirchlicher Trägerschaft (staatlich anerkannt).

Weiter gibt es insgesamt **223 „Fachhochschulen und Hochschulen ohne Promotionsrecht“**, davon 105 „Fachhochschulen und Hochschulen ohne Promotionsrecht“ in staatlicher Trägerschaft, 97 „Fachhochschulen und Hochschulen ohne Promotionsrecht“ in privater Trägerschaft (staatlich anerkannt) sowie 21 „Fachhochschulen und Hochschulen ohne Promotionsrecht“ in kirchlicher Trägerschaft (staatlich anerkannt).

Zahl der Studierenden insgesamt (2011/2012):¹

2.377.034 (davon weibl.: 1.123.607). Davon an

Universitäten	1.542.226
Päd. Hochschulen (B-W.)	23.221
Theol. HS	2.443
KunstHS	33.340
Fachhochschulen	744.150 (ohne Verw.FH)
Verw.FH	31.654
Fachhochschulen insges.	775.804

Die Zahl der Studierenden an Fachhochschulen hat also mittlerweile rd. 50% der Zahl der Studierenden an Universitäten erreicht.

Die größte Fachhochschule Deutschlands, die Fachhochschule Köln, umfasst 1.000 Mitarbeiter davon 600 wiss. Mitarbeiter, davon 430 Professoren sowie 19.700 Studenten (SS 2012) in 75 Studiengängen. Mit 44 Prozent stellen die Studierenden der Ingenieurwissenschaften den größten Anteil. Die Frauenquote liegt bei rund 35%.² Die zweitgrößte ist die Hochschule München. Sie umfasst fast 1.400 Mitarbeiter, davon ca. 500 Professoren sowie ca. 16.400 Studenten (WS 2011/12) (männliche: 10.763, weibliche: 5.676) in über 60 Studiengängen der Bereiche Natur-/Ingenieurwissenschaften, Wirtschaft, Sozialwesen und Design. Die Frauenquote beträgt ca. 35%.³

Die kleinsten Fachhochschulen umfassen ab 25 Studierenden (in Einzelstandorten), vor allem im privaten Bereich.⁴

¹ https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/StudierendeHochschulenVorb2110410128004.pdf?__blob=publicationFile, Aufruf: 06.08.2012

² http://de.wikipedia.org/wiki/Fachhochschule_K%C3%B6ln, Aufruf: 05.08.2012

³ http://de.wikipedia.org/wiki/Fachhochschule_M%C3%BCnchen, Aufruf: 05.08.2012

⁴ https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/StudierendeHochschulenVorb2110410128004.pdf?__blob=publicationFile, Aufruf: 05.08.2012

Die Studienerfolgsquoten an Universitäten und Fachhochschulen liegen 2010 in allen vergleichbaren Studiengängen in den Immatrikulationsjahrgängen 1999-2002 um jeweils 12 bis 14% besser an den Fachhochschulen als an Universitäten (schwankend an Universitäten zwischen 70,0 und 70,7 und bei den Fachhochschulen zwischen 82,3 und 84,3%).⁵

Noch deutlicher sind die Differenzen zwischen Universitäten (U) und Fachhochschulen (F) beim **Studienabbruch im Bachelor-Studium**, bezogen auf den Studienbeginn 2006/2007 (in Prozent):

Rechts-/Wirtschafts-/Sozialwesen 24 (U) zu 6 (F),
Mathematik/Naturwissenschaften 39 (U) zu 30 (F),
Agrar-/Forst-/Ernährungswissenschaften 33 (U) zu 21 (F)
und

Ingenieurwissenschaften 48 (U) zu 30 (F).⁶

Natürlich gibt es viele Gründe für Studienabbruch (zu schmale Informationsbasis bei der Fachwahl; falsches Bild vom Fach; zu große Differenz zwischen den attraktiv erscheinenden Realitäten in der Berufspraxis und der Selbstdarstellung und dem curricularen Aufbau als Studienfach, usw.). Aber die Differenzen zwischen den Hochschularten lassen sich damit nur unzureichend erklären. Die volkswirtschaftlichen Kosten und die individuellen Kosten des jeweiligen Scheiterns sind jedenfalls hoch.

⁵ https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/Erfolgsquoten5213001107004.pdf?__blob=publicationFile, Aufruf: 05.08.2012

⁶ Vgl. HIS-Magazin 3-2012, S. 3f.

Fachhochschulen auf Erfolgskurs Studentenzahl und Forschungsaktivitäten haben stark zugenommen / Schavan: „Internationale Strahlkraft wächst“

Fachhochschulen sind bei Studierenden attraktiver denn je. In den vergangenen zehn Jahren ist die Zahl der Studierenden an Fachhochschulen von 411.000 (Wintersemester 1999/2000) auf 684.000 (Wintersemester 2010/2011) gewachsen – also um 66%. Auch die Zahl der Fachhochschulen ist in diesem Zeitraum um ein Drittel gestiegen – von 152 auf 209. „Fachhochschulen bilden praxisnah aus, forschen anwendungsorientiert und kooperieren mit den Unternehmen in der Region – das macht sie attraktiver denn je“, lobte Bundesforschungsministerin Annette Schavan die starke Rolle der Fachhochschulen im deutschen Wissenschaftssystem.

Besonders erfreulich ist dabei, dass die Fachhochschulen immer stärker auf dem Gebiet der Forschung punkten. Alles in allem sind von 2006 bis 2011 rund 316 Millionen Euro in Forschung und Lehre von Fachhochschulen geflossen. So hat das BMBF seine Mittel für das Programm „Forschung an Fachhochschulen“, die unter der Vorgängerregierung im Jahr 2005 noch 10,5 Millionen Euro betragen, inzwischen nahezu vervierfacht. Damit wurden von 2006 bis 2011 allein in diesem Programm an 110 Fachhochschulen weit über 800 Forschungsvorhaben mit insgesamt 175 Millionen Euro gefördert. Zudem beteiligen sich Fachhochschulen immer erfolgreicher an anderen BMBF-Forschungsprogrammen. Diese umfassen eine große Spannweite – von den Fachprogrammen auf den Feldern „Neue Technologien“, „Lebenswissenschaften“ und „Klima, Umwelt, Energie“ über den Spitzencluster-Wettbewerb bis hin zur Innovationsförderung in den Neuen Ländern. „Diese Zahlen bestätigen uns in unserem Kurs, Forschungsaktivitäten an Fachhochschulen weiter zu fördern“, so Schavan.

Fachhochschulen machen neben starker Forschung auch durch innovative Lehrkonzepte von sich reden. Im Qualitätspakt Lehre wurden insgesamt 78 Projekte an Fachhochschulen ausgewählt. Das Ziel des Hochschulpaktes von Bund und Ländern, die Zahl der Studienplätze an Fachhochschulen zu erhöhen, wurde erreicht. Von den zwischen 2007 bis 2010 neu geschaffenen 185.000 Studienplätzen entfielen mehr als zwei Drittel auf Fachhochschulen. „Als Ausbildungsstätten für hochqualifizierten wissenschaftlichen Nachwuchs und als wichtige Partner für Unternehmen und Einrichtungen in der anwendungsorientierten Forschung sind die Fachhochschulen eine immer stärker werdende Säule unseres Wissenschaftssystems“, sagte die Ministerin. So stellen Fachhochschulen beispielsweise mehr als die Hälfte aller Ingenieurstudienplätze bundesweit.

Als Paradebeispiel für die zunehmende internationale Strahlkraft von Fachhochschulen nannte Schavan die Deutsch-Jordanische Hochschule in Amman. Sie ist nach dem Modell der deutschen Fachhochschulen mit praxisorientierten Studieninhalten aufgebaut. Konsortialführer auf deutscher Seite ist die Fachhochschule Magdeburg-Stendal. An der Deutsch-Jordanischen Hochschule sind 2.500 Studierende in insgesamt 26 Studiengängen eingeschrieben. Die Hochschule ist bereits jetzt stark vernetzt und arbeitet mit 60 Partnerinstitutionen zusammen.

Weitere Informationen finden Sie im Internet unter:
<http://www.bmbf.de/de/fh-foerderkarte.php>
sowie <http://www.bmbf.de/de/19918.php>.

Quelle: <http://idw-online.de/de/news493598>

Anzeige

Zwei postgraduale Studienprogramme für erweiterte berufliche Chancen

- *Diploma in Higher Education Development*
- *Diploma in Science Management (Organizational and Change Management)*

Die Förderung von Forschung, Lehre sowie ihrer Institutionen als Ihr Berufsfeld?

Wissenschaftsmanagement – Wissenschaftsförderung – Akademische und administrative Personal- und Organisationsentwicklung – fundierte Ausbildung von Moderator/innen und Berater/innen

Eine Aus- und Weiterbildungschance sowohl für

- Anfänger/innen, die eine attraktive und praktische Vorbereitung suchen, als auch für
- berufserfahrene Praktiker, die nachträgliche Grundlegung und Weiterbildung suchen.

Sie orientieren sich noch und erwägen Alternativen zur akademischen Karriere?

Forschung und Lehre sind als Berufsfeld verlockend.

Aber es ist unsicher, ob Sie als Wissenschaftler/in dort bleiben können.

Neben diesem Ziel sollten Sie sich daher rechtzeitig mit sehr überschaubarem Aufwand ein zweites Standbein aufbauen – eine zweite Existenz.

Vielleicht finden Sie es tatsächlich attraktiver, Wissenschaft (insbesondere Forschung und Lehre, Weiterbildung) zu ermöglichen, zu fördern, zu organisieren, Wandel zu begleiten – zu managen. Dabei reicht das Spektrum von der Steuerung wissenschaftlicher Einrichtungen bis zum professionellen Programmaufbau und der Moderation hochschulinterner Weiterbildung im Rahmen von Personal- und Organisationsentwicklung (mit der Variante "professioneller Hochschuldidaktik").

Sie sind an diesem Feld "Wissenschaftsförderung, Personal- und Organisationsentwicklung, Steuerung von wissenschaftlichen Einrichtungen" (i.w.S.) interessiert? Schon entschlossen, in diesem Feld neu Fuß zu fassen?

Variante A

Das IWBB bietet Ihnen dafür die beiden o.g., eng verknüpften Studienprogramme. Es handelt sich um ein berufsbegleitendes Studium im Umfang eines Masterstudiums, das durch die Anerkennung von Voraktivitäten reduziert werden kann. Auch Abschnitte werden zertifiziert.

Sie gewinnen damit in einem stürmisch wachsenden Arbeitsmarkt-Segment eine hervorragende Basis für eine haupt- oder nebenberufliche Tätigkeit, z.B. (zunächst) als Referent/in (o.ä.) in Hochschule und Wissenschafts(förder-)organisationen oder Moderator/in in hochschul-/wissenschaftsbezogener Weiterbildung.

Variante B

Sie sind bereits berufserfahren und haben Funktionen in der Organisation, Förderung, Unterstützung von Wissenschaft (im weiten Sinne: Wissenschaftsmanagement) übernommen?

Sie wollen sich weiter fundieren bzw. professionalisieren? Ihre Kompetenz erweitern?

Für über 30 einschlägige Funktionen in Hochschulen, Forschungszentren und Förderorganisationen gab es bisher keine oder erst wenige Ausbildungsmöglichkeiten – mit erheblichen Qualitätsdifferenzen. Nun gibt es Abhilfe.

Das IWBB bietet Ihnen zwei funktional angepasste Profile mit insgesamt 5 Spezialisierungsmöglichkeiten, die sich inhaltlich und zeitlich hochgradig individuell auf Ihre Bedürfnisse und Voraussetzungen einstellen lassen. Die berufsbegleitenden Programme sind nicht semestergebunden, können jederzeit beginnen. Leiter des Studienprogramms ist Prof. Dr. Wolff-Dietrich Webler.

Ist Ihr Anfangsinteresse geweckt?

Dann holen Sie sich nähere Informationen über unsere Webseite www.iwbb.de oder wenden sich direkt an das IWBB bzw. den Studiengangsleiter unter webler@iwbb.de.

IWBB, Bündlerstraße 1-3, 33613 Bielefeld.

**IWBB**

Institut für Wissenschafts- und Bildungsforschung Bielefeld
Bielefeld Institute for Research on Science and Education
 Forschen - Entwickeln - Begleiten - Beraten - Fortbilden

**Philipp Pohlenz:
Datenqualität als Schlüsselfrage der Qualitätssicherung von Lehre und Studium**



Hochschulen wandeln sich zunehmend zu Dienstleistungsunternehmen, die sich durch den Nachweis von Qualität und Exzellenz gegen ihre Wettbewerber durchsetzen müssen.

Zum Vergleich ihrer Leistungen werden verschiedene Evaluationsverfahren herangezogen. Diese stehen jedoch vielfach in der Kritik, bezüglich ihrer Eignung, Leistungen der Hochschulen adäquat abzubilden.

Verfahren der Evaluation von Lehre und Studium wird vorgeworfen, dass ihre Ergebnisse bspw. durch die Fehlinterpretation hochschulstatistischer Daten und durch die subjektive Färbung studentischer Qualitätsurteile verzerrt sind.

Im Zentrum des vorliegenden Bandes steht daher die Untersuchung von potenziellen Bedrohungen der Aussagefähigkeit von Evaluationsdaten als Steuerungsinstrument für das Management von Hochschulen.

*ISBN 3-937026-63-0, Bielefeld 2009,
170 Seiten, 22.80 Euro*

Bestellung - E-Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22

Liebe Leserinnen und Leser,

nicht nur in dieser lesenden Eigenschaft (und natürlich für künftige Abonnements) sind Sie uns willkommen. Wir begrüßen Sie im Spektrum von Forschungs- bis Erfahrungsberichten auch gerne als Autor/in. Wenn das Konzept der „Qualität in der Wissenschaft“ Sie anspricht - wovon wir natürlich überzeugt sind - dann freuen wir uns über Beiträge von Ihnen in den ständigen Sparten

- Qualitätsforschung,
- Qualitätsentwicklung/-politik,
- Anregungen für die Praxis/Erfahrungsberichte, aber ebenso
- Rezensionen,
- Tagungsberichte,
- Interviews.

Die Hinweise für Autorinnen und Autoren finden Sie unter: www.universitaetsverlagwebler.de.

Anzeigenannahme für die Zeitschrift „Qualität in der Wissenschaft“

Die Anzeigenpreise: auf Anfrage beim Verlag

Format der Anzeige: JPEG- oder EPS-Format, mindestens 300dpi Auflösung

UWW UniversitätsVerlagWebler, Der Fachverlag für Hochschulthemen, Bündler Straße 1-3 (Hofgebäude), 33613 Bielefeld, Fax: 0521 - 92 36 10-22

Kontakt: info@universitaetsverlagwebler.de

Hauptbeiträge der aktuellen Hefte Fo, HSW, HM, P-OE und ZBS

Auf unserer Homepage www.universitaetsverlagwebler.de erhalten Sie Einblick in das Editorial und Inhaltsverzeichnis aller bisher erschienenen Ausgaben.

Fo

Forschung

Politik - Strategie - Management

Forschung 1+2/2012
Brückenschlag im Dialog zwischen Wissenschaft und allgemeiner Öffentlichkeit

Wolff-Dietrich Webler
Brückenschlag zwischen Wissenschaft und allgemeiner Öffentlichkeit – das gesellschaftliche Umfeld

FO-Gespräch mit Markus Weißkopf, Wissenschaft im Dialog (WiD) mit Andrea Frank, Stifterverband

Überblick: Auftrag, Zielgruppen und Initiativen der WiD GmbH

Science-Festivals werden populärer – nur Stadtmarketing/Popularisierung/PR von Wissenschaft?
Als Beispiele: Die Science Festivals in Bielefeld („Geniale“) und Lübeck („Hanse trifft Humboldt“)

Stefan Stöcklin
„Citizen Science“ – Bürgerwissenschaft – soll Bürgerinnen und Bürger zu Forschenden machen
Die neuen Medien erleichtern die Mobilisierung

Peer Pasternack & Henning Schulze
Wissenschaftsräte zwischen Autonomie und Verbindlichkeit

Michael Krohn
Forschungsmanagement durch Großgruppenverfahren – das Beispiel World Café

Rezension:
Michael Bailey & Des Freedman:
The Assault on Universities:
A Manifesto for Resistance
(Christian Potschka)

HSW

Das Hochschulwesen

Forum für Hochschulforschung, -praxis und -politik

HSW 3/2012

Hochschulentwicklung/-politik

Thomas Günther, Sebastian John, Nancy Großmann, Ulrike Henke & Bianca Schönherr
Gestaltungsmöglichkeiten eines staatlichen Mittelverteilungssystems im Hochschulsektor:
Ein Ländervergleich

Johannes Balve
Quo vadis universitas?
Kritische Bemerkungen aus der Außenperspektive

Thea Rau, Andrea Kliemann, Jörg M. Fegert & Marc Allroggen
Schulung von Beratern an Hochschulen, Universitäten und Studentenwerken zum Umgang mit gefährdeten Studierenden

Hartmut Hinneberg
Der „Test für Medizinische Studiengänge“ im Auswahlverfahren der Hochschulen

Hochschulforschung

Philipp Budka, Claudia Schallert, Lukas Mitterauer & Markus Hintermayer
Neue Entwicklungen und Herausforderungen im sozialwissenschaftlichen Studium

Anregungen für die Praxis/
Erfahrungsberichte

Andreas D. Schulz
Mobiles Online-Projektseminar
„Internetpolitik und schwache Interessen“

HM

Hochschulmanagement

Zeitschrift für die Leitung, Entwicklung und Selbstverwaltung von Hochschulen und Wissenschaftseinrichtungen

HM 3/2012
Lehrcontrolling und -anreize – Tagung Performance Management im Hochschulbereich 2011 in Braunschweig

Entwicklung, Gestaltung und Verwaltung von Hochschulen und Wissenschaftseinrichtungen

Marcel Clermont, Britta Gerets & Matthias Meyer
Wie lässt sich Kompetenz an Hochschulen (v)ermitteln?
Eine Fragestellung für das Hochschulmanagement bzw. -controlling

Axel Oberschelp
Berichtswesen und Hochschul-Steuerung
Was und wie viel müssen Hochschul-Manager über Erfolg in der Lehre wissen?

Philipp Pohlentz, Olaf Ratzlaff & Markus Seyfried
Studiengang Fact Sheets für eine evidenzbasierte Steuerung von Lehre und Studium

Susanne In der Smitten & Michael Jaeger
Hochschulische Lehrleistungen im Kontext der W-Besoldung

P-OE

Personal- und Organisationsentwicklung in Einrichtungen der Lehre und Forschung

Ein Forum für Führungskräfte, Moderatoren, Trainer, Programm-Organisatoren

P-OE 2/2012
Aufbau einer Organisations- und Personalentwicklung mit indischen Hochschulpartnern

Personal- und Organisationsentwicklung/-politik

Birgit Barden
Organisations- und Personalentwicklung mit indischen Hochschulpartnern beim Aufbau von International Offices

Birgit Barden
Das Praktikum als Instrument der Personalentwicklung

Ricarda Mletzko
Workshopdesign im interkulturellen Kontext – Erfahrungen aus einem Organisations- und Personalentwicklungsprojekt mit indischen Hochschulen

Birgit Barden & Ricarda Mletzko
Auswertung des Projekts und Transferhinweise (Lessons learned)

Nathalie Baumeister, Anne-Marie Fleischmann, Katrin Klink, Karina Muntean & Anke Diez
Kollegiale Beratung als Instrument zur Reflexion und Erweiterung der Lehr- und Lernerfahrungen

PE-/OE-Forschung

Sabine Brendel, Sarah-Lena Krömker, Anja Riedel & Manuel Scholz
Wo Coaching eine „schmerzende Lücke“ füllt – Zur Einschätzung von Coaching als Instrument der akademischen Personalentwicklung in Hochschulen

ZBS

Zeitschrift für Beratung und Studium

Handlungsfelder, Praxisbeispiele und Lösungskonzepte

ZBS 3/2012
Fort- und Weiterbildung für Beratungspersonal an Hochschulen

Beratungsentwicklung/-politik

Elke Mittag
Aus- und Fortbildung niedersächsischer Studienberaterinnen und Studienberater durch die Zentrale Koordinierungsstelle für Studieninformation und -beratung in Niedersachsen

Martin Scholz
Der Weg zum GIBeT-Zertifikat – Das Umsetzungskonzept zum Fortbildungscurriculum der Gesellschaft für Information, Beratung und Therapie an Hochschulen

Sigrid Eicken, Oliver Orth & Irmgard Rieder
Die Professionalisierung der Studienberatung in Baden-Württemberg – wie aus einer Fortbildungsidee ein landesweit getragenes Programm wurde

Anregungen für die Praxis/ Erfahrungsberichte

Krischan Brandl
Das Weiterbildungsprogramm des csnd e.V. für Mitarbeiterinnen und Mitarbeiter der Career Services

Karin Gavin-Kramer
Qualifikations- und Eingruppierungs-Dumping bei der Studienberatung?

Thea Rau, Birgit Luderer & Claus Kaiser
Qualitätssicherung in der psychosozialen Beratung von Studierenden Einführung eines Qualitätsmanagementsystems im Studentenwerk Ulm



Für weitere Informationen

- zu unserem Zeitschriftenangebot,
- zum Abonnement einer Zeitschrift,
- zum Erwerb eines Einzelheftes,
- zum Erwerb eines anderen Verlagsproduktes,
- zur Einreichung eines Artikels,
- zu den Autorenhinweisen

oder sonstigen Fragen, besuchen Sie unsere Verlags-Homepage:

www.universitaetsverlagwebler.de

oder wenden Sie sich direkt an uns:

E-Mail:
info@universitaetsverlagwebler.de

Telefon:
0521/ 923 610-12

Fax:
0521/ 923 610-22

Postanschrift:
UniversitätsVerlagWebler
Bünder Straße 1-3
33613 Bielefeld

im UniversitätsVerlagWebler erhältlich:

Reihe Hochschulwesen: Wissenschaft und Praxis

**Wolff-Dietrich Webler (Hg.):
Studieneingangsphase?
Das Bachelor-Studium braucht eine neue Studieneingangsphase!**

Band I: Studierfähigkeit für ein frei(er)es Studium

Die Qualität der Studieneingangsphase hat wesentlichen - u.U. sogar entscheidenden - Einfluss auf den Studienerfolg. Trotzdem ist sie in der Vergangenheit in ihrer Bedeutung oft unterschätzt und vernachlässigt worden. Die Relevanz dieses Studienabschnitts wird jedoch inzwischen hoch eingestuft, eine Korrektur aufgrund z.T. falscher Signale aus der Bologna-Reform immer dringender. Für die Bewältigung des Übergangs in ein wissenschaftliches Studium aus der Schule oder aus dem Beruf ist wesentlich mehr nötig als ein Orientierungswochenende, (fach-)ein-führende Veranstaltungen und fachliche Begleittutorien. Mit der Feststellung „Das Bachelor-Studium braucht eine neue Studieneingangsphase - Studierfähigkeit für ein frei(er)es Studium!“ hatte das IWBB eine Initiative gestartet (u.a. eine Tagung), die den Erkenntnisstand zur Gestaltung der Studieneingangsphase bilanzieren, an solchen Fragen arbeitende Kolleg/innen zusammenführen und die Verbreitung besonders fortgeschrittener Modelle befördern sollte. Die Ergebnisse liegen nun in einem Doppelband vor.

Als zentrales Ziel wollen die Texte dazu beitragen, in den ersten beiden Semestern konzentriert die Studierenden zu befähigen, für sich anschließend ein motivierendes, durch Wahlmöglichkeiten möglichst selbst organisiertes, selbst verantwortetes und lerneffektives Studium zu organisieren.

Da das Themenfeld sehr groß ist, werden Fragen des Hochschulzugangs und von Auswahlverfahren nur am Rande angeschnitten. Die empirisch gewonnen Erkenntnisse zur Situation der Studierenden am Studienbeginn liegen bereits öffentlich vor, können also vorausgesetzt werden.

Im Mittelpunkt dieses Doppelbandes stehen daher in Band I die Ableitung und Begründung der Struktur des Problems und des Handlungsbedarfs (einschließlich des Beratungsbedarfs) sowie Gestaltungsprinzipien der Studieneingangsphase auf curricularer Ebene und der Ebene der Lehr-/ Lernprozesse. Behandelt wird die Frage, warum die Studieneingangsphase die ersten beiden Semester umfasst und wie sich Fachlichkeit und die Befriedigung des Orientierungsbedarfs über weite Strecken in den gleichen Veranstaltungen verschränken bzw. integrieren lassen, statt sie in Sonderveranstaltungen auszulagern.

Band II: Lösungsmodelle

Im Band II dieses Doppelbandes werden vielfältige Lösungen für Transfer und Gestaltungspraxis vorgelegt.

Diese Lösungen sind z.T. strategisch als umfassende Handlungsprogramme konzipiert, z.T. setzen sie bei einzelnen Lehrveranstaltungen an, z.T. bestehen sie aus einzelnen Lernmaterialien, die in besonderem Maße geeignet sind, einen Lernprozess in Richtung der Studienziele und der Ziele der Studieneingangsphase zu provozieren und zu unterstützen.

Das Spektrum der Beispiele wird durch besonders eindrucksvolle Modelle aus der Zeit vor der Bologna-Reform ergänzt, an denen auch heute noch viel gelernt werden kann.



ISBN 10: 3-937026-76-2, Bielefeld 2012, 477 S.,
Doppelband im Schuber, 69.50 Euro zzgl. Versandkosten

Bestellung - E-Mail: info@universitaetsverlagwebler.de, Fax: 0521/ 923 610-22