

Das Hochschulwesen

Forum für Hochschulforschung, -praxis und -politik

Institutional Research - die Hochschulen lernen mehr über sich selbst

- **Organisationswandel in Wissenschaft und Hochschulen - Kooperation - Vernetzung - Fusion**
HSW-Gespräch mit Dr. Josef Lange
- **Eine Studie über die Interrater-Reliabilität zur Prüfung beim Erstjahresstudium der Psychologie**
- **Einfluss von Lernverhalten und Lernstrategien auf die paradoxen Geschlechtsunterschiede in den Erfolgsraten im Wiener Medizinstudium**
- **Förderprogramm „Hochschulforschung als Beitrag zur Professionalisierung der Hochschullehre“ des Bundesministeriums für Bildung und Forschung**
- **Angenehm, konstruktiv – und nicht allzu wirkungsvoll? Lehrveranstaltungsevaluation aus der Sicht von Studierenden, Lehrenden und Evaluationsanbietern**

6 | 2007

Herausgeber

Christa Cremer-Renz, Prof. Dr. päd.,
bis Mai 2006 Präsidentin der Universität Lüneburg

Gustav-Wilhelm Bathke, Prof. Dr. sc.phil.,
Universität Halle-Wittenberg

Ludwig Huber, Prof. em. Dr. phil., Dr. h.c.,
Universität Bielefeld

Jürgen Lüthje, Dr. jur., Dr. h.c.,
bis Oktober 2006 Präsident der Universität Hamburg

Beate Meffert, Prof. Dr.-Ing.,
Humboldt-Universität zu Berlin

Klaus Palandt, Dr. jur., Min. Dirig. a.D.,
Landesbergen b. Hannover

Ulrich Teichler, Prof. Dr. phil.,
Universität Kassel

Wolff-Dietrich Webler, Prof. Dr. rer. soc.,
Institut für Wissenschafts- und Bildungsforschung
Bielefeld (geschäftsführend)

Andrä Wolter, Prof. Dr. phil., TU Dresden,
bis Dezember 2006 Hochschul-Informations-System
GmbH, Hannover

Herausgeber-Beirat

Hermann Avenarius, Prof. Dr., Frankfurt (M.)

Ralf Bartz, Univ. Kanzler, Hagen

Jost Bauer, Prof., Reutlingen

Winfried Benz, Dr., Gen. Sekr. WR i. R., Köln

Christian Bode, Dr., Gen. Sekr. DAAD, Bonn

Rüdiger vom Bruch, Prof. Dr., Berlin

Gertraude Buck-Bechler, Prof. em. Dr., Berlin

Matthias Bunge, Min.Dirig., Wiesbaden

Rik van den Bussche, Prof. Dr., Hamburg

Michael Deneke, Dr., Darmstadt

Gerhild Framhein, Dr., Konstanz

Karin Gavin-Kramer, M.A., Berlin

Gernot Graebner, akad. Dir. Dr., Deutsche Gesellschaft für
wissenschaftliche Weiterbildung und Fernstudium
(DGWF), Bielefeld

Lydia Hartwig, Dr., stellv. Leiterin, Bayer. Staatsinstitut für
Hochschulforschung und -planung

Jürgen Heß, Dr., Bonn

Sigurd Höllinger, Prof. Dr., Sektionschef im BM. Wiss. u.
Fo., Wien

Gerd Köhler, Frankfurt am Main

Artur Meier, Prof. Dr., Berlin

Sigrid Metz-Göckel, Prof. Dr., Dortmund

Jürgen Mittelstraß, Prof. Dr., Konstanz

Ronald Mönch, Prof. Dr. h.c., Bremen

Jan H. Olbertz, Prof. Dr. sc., Halle, Kultusminister des
Landes Sachsen-Anhalt

Jürgen Schlegel, Min.Dirig., Gen. Sekr. BLK, Bonn

Klaus Schnitzer, Dr., Hannover

Carl-Hellmut Wagemann, Prof. em. Dr.-Ing., Berlin

Karl Weber, Prof. Dr., Bern

Johannes Wildt, Prof. Dr. Dr. h.c., Dortmund; Bundesvorsit-
zender der Arbeitsgemeinschaft für Hochschuldidaktik
(AHD)

Hinweise für die Autoren

Senden Sie bitte zwei Exemplare des Manuskripts in
Papierform sowie einmal in Dateiform (kann als Daten-CD
der Papierform beigelegt oder per E-Mail zugeschickt wer-
den) an die Redaktion (Adresse siehe Impressum). Beiträge
werden nur dann angenommen, wenn die Autor/innen
den Gegenstand nicht gleichzeitig in einer anderen Zeit-
schrift behandeln.

**Wichtige Vorgaben zu Textformatierungen und beigefüg-
ten Fotos, Zeichnungen sowie Abbildungen erhalten Sie in
den „Autorenhinweisen“ auf unserer Verlags-Homepage
„www.universitaetsverlagwebler.de“.**
Ausführliche Informationen zu den in diesem Heft aufge-
führten Verlagsprodukten erhalten Sie ebenfalls auf der
zuvor genannten Verlags-Homepage.

Impressum

Verlag und Abonnementverwaltung

UVW UniversitätsVerlagWebler
Der Fachverlag für Hochschulthemen
Bünder Str. 1-3
33613 Bielefeld

Tel.: (0521) 92 36 10-12, Fax: (0521) 92 36 10-22

Satz:

K. Gerber, E-Mail: gerber@universitaetsverlagwebler.de
Übersetzung editorial: Jessica Kleinehelftwes

Druck:

Hans Gieselmann,
Ackerstr. 54, 33649 Bielefeld

Anzeigen:

Das HSW veröffentlicht Verlagsanzeigen, Ausschreibungen und
Stellenanzeigen. Aufträge sind an den Verlag zu richten. Die jeweils
gültigen Anzeigenpreise sind folgender Homepage zu entnehmen:
„www.hochschulwesen.info“.

Erscheinungsweise: 6mal jährlich

Redaktionsschluss: 14. Dezember 2007

Bezugspreis:

Jahresabonnement 80 Euro/156 SFR, Einzelpreis 13.50 Euro
Alle Preise verstehen sich zuzüglich Versandkosten.

Das Jahresabonnement verlängert sich automatisch um 1 Jahr,
wenn es nicht 6 Wochen vor Jahresende gekündigt wird.

Grafik:

Ute Weber Grafik Design, München
Gesetzt in der Linotype Syntax Regular

Copyright: UVW UniversitätsVerlagWebler

Die mit Verfassernamen gekennzeichneten Beiträge geben nicht in
jedem Falle die Auffassung der Herausgeber bzw. Redaktion wieder.
Für unverlangt eingesandte Manuskripte/Rezensionsexemplare
wird keine Verpflichtung zur Veröffentlichung/Besprechung über-
nommen. Sie können nur zurückgegeben werden, wenn ausrei-
chendes Rückporto beigefügt ist. Der Nachdruck von Artikeln, auch
auszugsweise, sowie die Verwendung für Rundfunk/Fernsehen ist
nur mit Quellenangabe und Genehmigung des Verfassers gestattet.

Das Hochschulwesen

Forum für Hochschulforschung, -praxis und -politik

Editorial

161

HSW-Gespräche

Das Hochschulwesen
Organisationswandel in Wissenschaft und Hochschulen - Kooperation - Vernetzung - Fusion
HSW-Gespräch mit Dr. Josef Lange,
Staatssekretär im niedersächsischen Ministerium
für Wissenschaft und Kultur

162

Hochschulforschung

Arild Raaheim
Eine Studie über die Interrater-Reliabilität zur Prüfung beim Erstjahresstudium der Psychologie

167

Oskar Frischenschlager, Lukas Mitterauer, Irina Scharinger & Gerald Haidinger
Einfluss von Lernverhalten und Lernstrategien auf die paradoxen Geschlechtsunterschiede in den Erfolgsraten im Wiener Medizinstudium

175

Dokumentation

Förderprogramm „Hochschulforschung als Beitrag zur Professionalisierung der Hochschullehre“
des Bundesministeriums für Bildung und Forschung

180

Hochschulentwicklung/-politik

Boris Schmidt
Angenehm, konstruktiv – und nicht allzu wirkungsvoll?
Lehrveranstaltungsevaluation aus der Sicht von Studierenden, Lehrenden und Evaluationsanbietern

183

Rezension

Uwe Kühler: Interkulturelle Hochschullehre: Internationalisierung am Beispiel der Amerikanistik.
(*Judith Ricken & Bianca Roters*)

191

Tagungsbericht

Exzellenz auch in der Lehre?
Erste Bilanztagung zur Exzellenzinitiative
(*René Krempkow*)

192

Seitenblick auf die Schwesterzeitschriften

Hauptbeiträge der parallelen Hefte IVI, P-OE, QiW, HM und ZBS

IV

Anke Hanft (Hg.) Grundbegriffe des Hochschulmanagements

Das Buch liefert grundlegende Informationen zu Managementkonzepten und -methoden sowie zu den derzeit diskutierten Reformansätzen im Hochschulbereich.

Erstmalig werden dabei auch die durch den Einsatz der Informations- und Kommunikationstechnologien in Lehre und Administration ausgelösten Veränderungen umfassend berücksichtigt.

Etwa 100 Begriffe werden in alphabetischer Reihenfolge erläutert.

Durch vielfältige Querverweise und ein umfassendes Stichwortverzeichnis ist sichergestellt, dass der Leserschnell und gezielt auf die ihn interessierenden Informationen zugreifen kann.

ISBN 3-937026-17-7,
Bielefeld 2004, 2. Auflage, 525 Seiten, 34.20



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

Sehr geehrte Abonentin, sehr geehrter Abonnent der Zeitschrift „Das Hochschulwesen“,

seit nunmehr 11 Jahren ist der Preis für „Das Hochschulwesen“ stabil geblieben.

In dieser Zeit gab es erhebliche Veränderungen.

Das HSW wurde 2001 vom eigens dafür gegründeten UniversitätsVerlagWebler "aufgefangen", als der Luchterhand-Verlag (Wolters & Kluiver, Amsterdam) sich aus allen Hochschulthemen zurückzog. Trotz des Aufbaus unseres völlig neuen Verlages haben wir den Preis gegen alle Schwierigkeiten stabil gehalten, dabei den Umfang pro Heft sogar von 36 auf 44 Seiten steigern können. Jetzt müssen wir aber der Kostenentwicklung folgen, auch, um den Service zu verbessern, mehr Gestaltungsspielraum zu gewinnen und Ihnen eine noch attraktivere Zeitschrift bieten zu können.

Dabei wollen wir uns auf die Mindestanhebung beschränken, die aber nach 11 Jahren Abstinenz sich nicht in Cent-Beträgen erschöpfen kann.

Ab 2008 setzen wir für das Jahres-Abonnement pro Heft 2 Euro mehr, also statt bisher 80 Euro nun 92 Euro ein.

Im Vergleich zu anderen Fachzeitschriften ist dieser Preis noch immer sehr moderat.

Die Vorteile der Mehrfach-Abonnements unserer Zeitschriften (diese plus ihre Schwesterzeitschriften) bleiben auf Basis der neuen Preise bestehen. Wir planen, sie nun einige Zeit stabil halten zu können.

Wir hoffen und Ihr Verständnis und verbleiben mit freundlichen Grüßen

Ihr UniversitätsVerlagWebler

In diesem Heft gilt als Schwerpunkt der Beiträge: Langsam ist mehr Institutional Research anzutreffen - die Hochschulen lernen mehr über sich selbst. Diesen Eindruck vermitteln einige der nachfolgenden Aufsätze und das neu beschriebene Förderprogramm Hochschulforschung des BMFT, das hier dokumentiert wird. Es hat Jahre gedauert, aber allmählich wächst die Bereitschaft in den Hochschulen, ihre Methoden auch auf sich selbst anzuwenden.

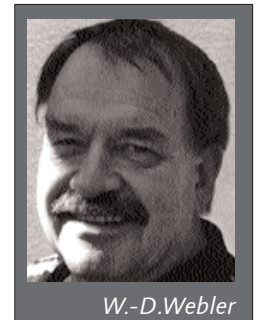
Dr. Josef Lange, Staatssekretär im niedersächsischen Ministerium für Wissenschaft und Kultur, entwickelt im **HSW-Gespräch über Organisationswandel in Wissenschaft und Hochschulen. Kooperation - Vernetzung - Fusion** Antworten auf die gegenwärtigen, ziemlich zeitgleichen großen Bewegungen, die in verschiedenen Regionen auf Bündnisse, wenn nicht sogar Zusammenschlüsse zu laufen.

Seite 162

Das Phänomen mangelnder Reliabilität der Notengebung bei Prüfungen ist zwar grundsätzlich bekannt, aber relativ unbearbeitet. **Arild Raaheim**, Professor für Psychologie an der Universität in Bergen (Norwegen) legt eine Studie über die **Interrater-Reliabilität zur Prüfung beim Erstjahresstudium der Psychologie** vor. An diesem gründlich gearbeiteten Beispiel wird gezeigt, wie wenig zuverlässig die Notengebung tatsächlich ist. Von ihren Noten in feinsten Abstufungen in Dezimalstellen weittragende Karriere-Entscheidungen abhängig zu machen, ist von den Leistungen der Prüfer her in keiner Weise gerechtfertigt. Das hatte man für Noten aus mündlichen Leistungen eher erwartet. Hier schriftliche Prüfungen mit gleichem Ergebnis vorzufinden, ernüchert die üblichen Erwartungen an Prüfungssysteme.

Seite 167

Das HSW veröffentlicht hier eine empirische Studie von **Oskar Frischenschlager, Lukas Mitterauer, Irina Scharinger & Gerald Haidinger** zum **Einfluss von Lernverhalten und Lernstrategien auf die paradoxen Geschlechtsunterschiede**



W.-D. Webler

in den Erfolgsraten im Wiener Medizinstudium. Während üblicherweise Frauen in Lernprozessen (bzw. deren Prüfung) besser abschneiden als Männer, ist dies im Wiener Medizinstudium umgekehrt. Diese paradoxe Situation ist nicht leicht aufzuklären. Weder unterschiedliches Lernverhalten noch Lernstrategien erklären sie. Der Beitrag bietet Erklärungen an, wirft aber auch neue Fragen auf.

Seite 175

Eine Zeitschrift für Hochschulforschung dokumentiert natürlich bereitwillig das **Förderprogramm Hochschulforschung des BMFT**, das zur weiteren Erforschung auch der Lehre bzw. Lehrleistungen an Hochschulen aufruft. Es ist sehr lange her, dass zunächst - in den frühen 70er Jahren - die VW-Stiftung und wenig später dann die DFG Förder-schwerpunkte in diesem Themengebiet eingerichtet hatten.

Seite 180

Boris Schmidt fragt in seinem vergleichenden Artikel: **Angenehm, konstruktiv - und nicht allzu wirkungsvoll? Lehrveranstaltungsevaluation aus der Sicht von Studierenden, Lehrenden und Evaluationsanbietern** nach der methodischen Anlage, aber vor allem der Wirkung von Lehrveranstaltungsevaluationen.

Seite 183

W.W.

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 Autorin und Autor.

Wenn das Konzept dieser Zeitschrift Sie anspricht - wovon wir natürlich überzeugt sind - dann freuen wir uns über Beiträge von Ihnen in den ständigen Sparten

- „Hochschulorschung“,
- „Hochschulentwicklung/-politik“,
- „Anregungen für die Praxis/Erfahrungsberichte“,
- aber ebenso „Rezensionen“,
- „Tagungsberichte“
- sowie „Interviews“.

Die Autorenhinweise finden Sie auf unserer Verlags-Homepage „www.universitaetsverlagwebler.de“.

Das Hochschulwesen



Josef Lange

Organisationswandel in Wissenschaft und Hochschulen Kooperation - Vernetzung - Fusion HSW-Gespräch mit Dr. Josef Lange, Staatssekretär im niedersächsischen Ministerium für Wissenschaft und Kultur

HSW: Herr Lange, seit geraumer Zeit häufen sich Kooperationen, Vernetzungen, ja Fusionen zwischen Einrichtungen im Bereich von Hochschule und Forschung. Noch vor kurzem hätte man derartige Verbindungen und Bündnisse unter diesen Partnern für unmöglich gehalten. Immer mehr Universitäten und Fachhochschulen kreieren gemeinsame Studiengänge, entwickeln gemeinsame Promotionsmöglichkeiten; in mehreren Regionen in Deutschland werden enge Kooperationen mit der Perspektive sogar der Fusion zwischen Universitäten und Max-Planck-Instituten, mit Fraunhofer-Instituten oder mit den Forschungszentren Jülich oder Karlsruhe vorbereitet. In Niedersachsen wird eine enge Vernetzung der TU Braunschweig, der TU Clausthal und der Universität Hannover betrieben. Vor wenigen Jahren fand die Fusion zwischen der Universität Lüneburg und der Fachhochschule Nordostniedersachsen mit Sitz in Lüneburg statt. Diese Vernetzungen und Fusionen in Forschung und Studium nützen offensichtlich beiden Seiten gleichermaßen.

1. Anlass, Ziele und erwarteter Nutzen

HSW: Auch wenn die Fälle unterschiedlich gelagert sind - was löst diese Betriebsamkeit aus?

Josef Lange: Die rasante Entwicklung der Wissenschaft in allen ihren Zweigen weltweit ist eine der größten Herausforderungen für die Hochschulen. Gegenwärtig - so betont der Präsident der Alexander von Humboldt-Stiftung, Professor Dr. Wolfgang Frühwald, immer wieder - sind mehr Menschen in der Wissenschaft tätig als in der Geschichte der Menschheit bisher zusammen. Daraus folgt, dass der internationale Wettbewerb in der Wissenschaft immer härter wird. Deshalb ist es bereits seit geraumer Zeit geboten, in den Hochschulen insbesondere bei Berufungen auf Professuren eine deutliche Profilierung durch Schwerpunktsetzung anzustreben. Keine Universität kann in allen in ihr vertretenen Fächern weltweit führend sein - das scheitert an der Verfügbarkeit exzellenten Personals und an den dafür erforderlichen finanziellen Ressourcen. Deshalb ergibt sich die Notwendigkeit - und die Exzellenzinitiative des Bundes und der Länder hat gezeigt, dass das an viel mehr Orten als erwartet möglich ist - , durch Vernetzungen und entscheidungsfähige, Institutionen übergreifende (Entscheidungs-) Strukturen wissenschaftlich „kritische Masse“ zu schaffen,

During the HSW-Talk **The change of organisation in Science and at Universities. Cooperation - Cross-linking - Fusion**, Dr. Josef Lange, permanent secretary in the Ministry for Science and Culture of Lower Saxony, is giving answers concerning the actual movements, which lead to confederacies in some regions and maybe even to mergers.

um im internationalen Wettbewerb konkurrenzfähig zu bleiben oder zu werden.

HSW: In welcher Lage befinden sich die jeweiligen Partner? Sind

- Anlass,
- Ziele und
- erwarteter Nutzen in allen Fällen gleich?

Nehmen wir als Beispiele mal die spektakuläre Fusion zwischen der Universität Lüneburg und der Fachhochschule Nordostniedersachsen einerseits und den Zusammenschluss von Forschungszentren und Universitäten, wie in Karlsruhe, Berlin, München oder Aachen andererseits. Kann man eine Typisierung solcher Bestrebungen benennen?

Josef Lange: Anlass und Ziele sind unterschiedlich. Bei der Fusion in Lüneburg ging und geht es darum, durch Fusion zweier von Größe, Leistungsfähigkeit und Finanzierung her an der Untergrenze tätige Hochschulen zu einer zumindest national und hoffentlich in einigen Feldern auch international konkurrenzfähigen Universität zu entwickeln, die den Bologna-Prozess und die daraus folgende hochschulinterne Differenzierung entsprechend der unterschiedlichen internen Leistungsfähigkeit ernst nimmt. In Karlsruhe wird eine Fusion von Universität und Helmholtz-Zentrum angestrebt: beide Einrichtungen zeichnen sich dadurch aus, dass bereits seit langer Zeit Institute in Universität und Zentrum gemeinsam von einer Forscherpersönlichkeit geleitet werden und insofern die Kooperation enger ist als an anderen Standorten. Gleiches gilt in bestimmten Bereichen zwischen der RWTH Aachen und dem Forschungszentrum Jülich. Hier geht es in beiden Fällen um das Vorstoßen in die weltweite Spitzengruppe technischer Universitäten. Ähnliches gilt für den Wissenschaftsstandort Göttingen, wo das in der Exzellenzinitiative erfolgreiche Zukunftskonzept gegründet

wird auf einen Göttingen Research Council, in dem Universität, Max-Planck-Institute und Leibniz-Institute gemeinsam die Strukturfragen des Standorts entscheiden.

HSW: Welche Fälle sind Ihrer Kenntnis nach in Planung bzw. bereits begonnen? War die Aufzählung oben schon vollständig?

Josef Lange: Auch im Ruhrgebiet sind die Hochschulen auf dem Weg zu weiterer Vernetzung. Zu nennen sind auch die Bestrebungen des Berliner Wissenschaftssenators, leistungsstarke Bereiche der drei Berliner Universitäten zusammenzuschließen. In Niedersachsen starten die TU Clausthal und die FH Braunschweig-Wolfenbüttel einen gemeinsamen Master-Studiengang. Außerdem arbeiten die drei technisch orientierten Universitäten Braunschweig, Clausthal und Hannover gemeinsam mit dem Nds. Ministerium für Wissenschaft und Kultur an Eckpunkten und einem Gesetzentwurf zur Schaffung einer Niedersächsischen Technischen Hochschule (NTH), der zu Beginn der kommenden Legislaturperiode in den Niedersächsischen Landtag eingebracht werden soll. Ich gehe davon aus, dass die institutionellen Vernetzungen in den kommenden Jahren weiter zunehmen werden.

HSW: Wie reagieren die Mutterorganisationen darauf - etwa bei Max-Planck- oder Fraunhofer-Instituten?

Josef Lange: Vernetzungen über gemeinsame Institutsleitungen sind bei der Fraunhofer-Gesellschaft vielfach Normalität und es gehört zu den Zielen der FhG, ihre Institute mit den Hochschulen auf diesem Weg zu verklammern. Auch die MPG strebt seit geraumer Zeit engere Verknüpfungen mit den Universitäten an. Aufgrund der Vereinbarung zwischen Max-Planck-Gesellschaft und Hochschulrektorenkonferenz über die Errichtung von International Max Planck Research Schools at Universities von 1999 sind inzwischen 49 IMPRS eingerichtet worden. Max-Planck-Institute sind an vielen Graduiertenkollegs, Sonderforschungsbereichen, Graduiertenschulen und Exzellenzclustern beteiligt, so dass auch hier die Hürden für Zusammenarbeit und Vernetzung bei Beibehaltung und gegenseitigen Achtung der Eigenständigkeit deutlich gesenkt wurden.

HSW: Wie steht der Staat, wie stehen die Länder solchen Bestrebungen gegenüber? Verschieben sich hier nicht auch Gewichte in der Forschungslandschaft der Bundesrepublik, die den Ländern nicht gleichgültig sein können? Können die Reaktionen der Länder mit Verhinderung, Duldung, Förderung oder sogar Flucht nach vorn bezeichnet werden?

Josef Lange: Forschungsförderung ist auch nach der Föderalismusreform I eine Gemeinschaftsaufgabe des Bundes und der Länder. Außerdem darf nicht vergessen werden, dass die Forschungsleistungen der Hochschulen in Deutschland immer noch größer sind als die aller außeruniversitären, öffentlich finanzierten Forschungseinrichtungen zusammen. Das gilt auch für die Finanzierung. Außerdem wartet die internationale Konkurrenz nicht darauf, dass Deutschland seine Forschungsstrukturen nach Modellen am „Grünen Tisch“ organisiert. Die historisch gewachsene institutionelle Differenzierung in der Forschung in Deutschland, die bis in das 19. Jahrhundert zurückgeht, muss zukunftsorientiert weiterentwickelt werden. Auf dem Weg zu dieser Entwick-

lung ist in den letzten Jahren viel geschehen und noch mehr in Arbeit, wie die Beispiele zeigen.

HSW: Wenn ein großer Forschungspartner bisher mit Bundesmitteln gefördert wurde; wie sieht die Zukunft der Finanzierung aus, wenn ein solches Institut in eine Fusion eintritt? Erzwingt die Weiterfinanzierung auch die Separierung der Haushalte und verhindert den letzten Schritt der Fusion, um größeren Einfluss des Bundes auf Universitäten zu verhindern?

Josef Lange: Wenn man zu klugen institutionellen Lösungen kommt, werden sich auch kluge Finanzierungsregelungen finden lassen. Die außeruniversitären Forschungseinrichtungen leben seit langem gut mit unterschiedlichen Partnern und unterschiedlichen Schlüssel der gemeinsamen Finanzierung. Hier lässt sich nur zu pragmatischen Einzelfalllösungen raten. Wenn zunächst Grundsatzfragen diskutiert und verfassungsgerichtsfest juristisch definiert werden sollen, dann läuft uns in der internationalen Konkurrenz die Zeit davon. Wir sind auch institutionell auf dem Weg zu Neuem und werden – so hoffe ich – dadurch dem Anspruch gerecht, die Rahmenbedingungen für „Wissenschaft als Erwartung der Unerwarteten“ angemessen zu gestalten.

2. Forschung

HSW: Trennen wir einmal die Vorgänge im Bereich der Studiengänge und in der Forschung und betrachten zunächst die Vorgänge im Bereich der Forschung. Hier scheint es großen Forschungszentren nicht zuletzt um internationale Sichtbarkeit in Rankings zu gehen, in die nur Universitäten einbezogen werden (etwa das Shanghai-Ranking). Sind die jeweils verfolgten Kooperations- und Fusionsziele richtig? Welche sind es im einzelnen?

Josef Lange: Der Blick auf die Rankings ist das Eine – bezogen auf Institutionen, aber in der Realität der Wissenschaft ist Reputation doch in der Regel eher auf das jeweilige Fach oder eine Fächergruppe bezogen. Insofern verschleiern die Rankings und die öffentliche Diskussion über sie außerhalb der Wissenschaft die Tatsache, dass auch die weltweit führenden Universitäten oder Forschungseinrichtungen nicht in allen Fächern die absolute Spitze bilden. Allerdings darf auch nicht verschwiegen werden, dass die regelmäßige Publikation von Rankings für Öffentlichkeit und Politik verführerisch wirkt, weil die differenzierte Darstellung von Entwicklungsprozessen sich hinter einer Tabelle verbirgt, mit der sich in der Mediengesellschaft öffentlichkeitswirksam trefflich argumentieren und streiten lässt.

HSW: Wollen die großen Forschungsinstitute bzw. -zentren in der Kooperation mit Universitäten u.a. ihren Zugang zu Promotionen erleichtern? Wollen sie besten Nachwuchs an sich binden?

Josef Lange: Bei der Promotion geht es vorrangig um Zusammenarbeit im Hinblick auf das Promotionsrecht der Universitäten. Hier haben sich in der Vergangenheit manche Fächer und manche Hochschulen gegenüber ihren Partnern aus außeruniversitären Forschungseinrichtungen gelegentlich wenig kooperativ gezeigt. Das muss Vergangenheit sein.

HSW: Wie steht es mit der schon länger praktizierten gemeinsamen Berufung in Professuren, die gleichzeitig an Universitäten und Forschungsinstituten angesiedelt sind? Die Professuren haben ja dann überwiegend Forschungsaufgaben. Sind die Regelungen zum Lehrdeputat beispielsweise ausreichend flexibel, um solche Aufgaben zuzulassen?

Josef Lange: Die Regelungen zur Wahrnehmung der Lehre sind hinreichend flexibel bei den unterschiedlichen Modellen gemeinsamer Berufung (Jülicher, Karlsruher, Berliner etc. Modell). Niedersachsen hat die Entscheidungskompetenz darüber in der Neufassung der Lehrverpflichtungsordnung den Hochschulleitungen übertragen. Damit müssten fast alle denkbaren Lösungen je nach örtlichen und institutionellen Bedürfnissen regelbar sein.

HSW: Wie steht es mit Kooperationen zwischen Universitäten und Fachhochschulen in der Forschung? Etwa auch beim Zugang von hervorragenden Absolventen der Fachhochschulen zur Promotion? Auch die Absolventen von Master-Studiengängen an Fachhochschulen sind berechtigt, zur Promotion zugelassen zu werden.

Josef Lange: In diesem Feld besteht noch erheblicher Verbesserungsbedarf, obwohl es bereits ein gemeinsames DFG-Graduiertenmodell einer Universität und einer Fachhochschule gab. An vielen Universitäten herrscht bei der Frage der Zulassung von Fachhochschulabsolventinnen und -absolventen zur Promotion noch mehr institutioneller Dünkel, denn Orientierung an den Leistungen und der Leistungsfähigkeit der Kandidatinnen und Kandidaten vor. Eine solche Haltung ist nicht gerechtfertigt – hier liegen die Aufgaben eindeutig bei den Universitäten, auch wenn eine Reihe Universitäten bereits gute Modelle kooperativer Promotion entwickelt hat und praktiziert. Auch zu dieser Frage empfiehlt sich die realitätsbezogene Lösung des Einzelfalls der Kooperation zwischen definierten Hochschulen statt abstrakt ideologisch mit den Vorurteilen der Vergangenheit zu argumentieren.

HSW: Bisher sind Dissertationen i.d.R. Beiträge zur Grundlagenforschung. Angesichts der immer stärker verwischten Grenzen zwischen grundlagen- und anwendungsbezogener Forschung und der Diskussion über Verwertungsgesichtspunkte wird auch über stärker anwendungsbezogene Dissertationen bis hin zu einer Sichtbarkeit dieser Ausrichtung in der Bezeichnung des Doktorgrades diskutiert. Wie stehen Sie zu dieser Entwicklung?

Josef Lange: Die Promotion dokumentiert die Fähigkeit zu eigenständiger wissenschaftlicher Leistung mit neuen Erkenntnissen. Ich halte nichts von einem „Berufsdoktorat“. Die Themen von Dissertationen in vielen Fächern zeigen, dass die Unterscheidung zwischen Grundlagen- und angewandter Forschung in vielen Bereichen obsolet wird. Jedenfalls ist dies kein Grund für eine Einführung eines neuen Abschlusses oder Titels. Hier kann man nur vor einer Grundsatzdiskussion mit unnötigen neuen Abgrenzungstendenzen zwischen den Hochschulen warnen. Eine solche Lösung diene weder der Forschung, noch den Forschenden noch den Institutionen und schon gar nicht dem internationalen Ansehen und der internationalen Reputation deutscher Hochschulen

HSW: Sehen Sie angesichts der nicht eben seltenen Diskriminierungen von FH-Absolventen ein eigenes Promotionsrecht der Fachhochschulen herannahen? Immerhin gilt innerhalb von Universitäten schon lange, dass derjenige prüfungsberechtigt ist, wer die entsprechende Prüfung selbst abgelegt hat - und Professoren an Fachhochschulen müssen in Deutschland (etwa im Gegensatz zur Schweiz) promoviert sein; es gibt mittlerweile sogar FH-Fachbereiche, in denen ein Drittel des Lehrkörpers habilitiert ist.

Josef Lange: Der Weg von Fachhochschulen hin zu einem Promotionsrecht würde diese nur in einen institutionellen Wettbewerb mit den Universitäten führen, den die Fachhochschulen angesichts der Finanzierungsmöglichkeiten der öffentlichen Hand, aber auch Privater nur verlieren können. Der Weg der Wahl ist die kooperative Promotion. Gleiches gilt für die Kooperation zwischen Universitäten und außeruniversitären Einrichtungen: das Promotionsrecht von Akademien und Forschungsinstituten gehört seit dem Ende der Teilung der Welt zwischen Ost und West zur Geschichte.

HSW: Sind die jeweils geplanten Kooperationen, Vernetzungen, ja Fusionen überhaupt das richtige Instrument - bezogen auf die verfolgten Ziele?

Josef Lange: Unterschiedliche Ziele erfordern unterschiedliche Wege. Für die internationale Konkurrenzfähigkeit brauchen wir zugleich größere und flexiblere Einheiten. Sie zu erreichen, erfordert jeweils individuelle Schritte. Ich glaube nicht, dass es fertige Rezepte für solche Prozesse gibt.

HSW: Sind die eingeschlagenen Wege und die Etappen in den Entwicklungsprozessen im Licht eines Change Management aus Ihrer Sicht richtig gewählt?

Josef Lange: Wege und Etappen sind unterschiedlich gewählt. Um beurteilen zu können, ob die Wege und Schritte richtig gewählt sind, müssen die Einzelfälle auch in ihren internen Entwicklungsschritten besser bekannt sein als sie es heute sind. Hinzu kommt, dass komplizierte Vernetzungs- oder gar Fusionsprozesse so gestaltet werden müssen, dass die beteiligten Handelnden in den Institutionen auf diesem Weg mitgenommen werden. Dazu ist zu viel Öffentlichkeit manchmal von Nachteil, weil vor der Öffentlichkeit immer auch die „Gesichtswahrung“ eine Rolle spielt, weil viel zu schnell die Frage nach Gewinnern und Verlieren im Fusionsprozess gestellt wird. Bezogen auf die Wissenschaft kann das Ziel nur lauten, alle müssen gewinnen (können).

HSW: Sind - soweit absehbar - die richtigen Akteure beteiligt? Fehlt eine Seite?

Josef Lange: Soweit absehbar sind die richtigen Akteure beteiligt. Wann es sinnvoll ist, externe Moderatoren oder Mediatoren einzubeziehen, kann nur im Einzelfall entschieden werden.

HSW: Wie sehen Sie die Rolle des Staates in diesen Prozessen?

Josef Lange: Der Staat kann angesichts der schnellen Entwicklung in der Wissenschaft nur Rahmenbedingungen und Finanzierung vorgeben und bereitstellen. Das bedeutet indes in einem staatlich gegründeten, staatlich organisierten und staatlich finanzierten Hochschul- und Forschungs-

system auch, dass die Ziel der Hochschulentwicklung als politische Ziele zu definieren sind und dann Politik und Administration an der Umsetzung mitzuwirken haben. Wie das im Einzelnen geschieht, hängt von den jeweiligen Strukturen und den handelnden Personen ab.

HSW: Was wird staatlicherseits an Unterstützung gebraucht :

- politisch (auch symbolisch-verbale Unterstützung)?
- rechtlich (Änderung von Gesetzen, begleitende oder grundlegende Verordnungen)?
- finanziell?

Josef Lange: Neben angemessenen rechtlichen und finanziellen Rahmenbedingungen benötigen Hochschulen und Forschungseinrichtungen über die Definition von Zielen hinaus auch verbale und symbolische Unterstützung. Fusionsprozesse sind immer anstrengend. Sie führen ins Unbekannte und ggf. ins Neue, in bisher nirgendwo Realisiertes. Insofern sind Ängste bei den Betroffenen nahe liegend. Deshalb gehören Ermutigung und Unterstützung dazu, aber auch Wertschätzung wissenschaftlicher und organisatorischer Leistungen in der Öffentlichkeit, denn auch hier geht es um Reputation, die schwer errungen und leicht verspielt ist. Auch die Hochschul- und Forschungspolitik bedarf symbolischer Handlungen, die Notwendigkeit von „symbolic politics“ gilt auch hier.

3. Bereich der Studiengänge

HSW: Nehmen wir nun die Kooperationen im Bereich der Studiengänge. Hier sind Kooperationen zwischen Universitäten und zwischen Universitäten und Fachhochschulen entstanden. Letztere galten lange Zeit als Tabu, obwohl der erste gemeinsame Studiengang - ein Diplom-Studiengang Medizintechnik von Universität Heidelberg und gerade gegründeter Fachhochschule Heilbronn - bereits 1972 entstanden war (bei voller Prüfungsberechtigung der Heilbronner Lehrenden). Häufiger sind heute Kooperationen im Medienbereich. Können Sie einige Beispiele nennen? Sind die jeweils verfolgten Ziele richtig? Welche sind es im einzelnen?

Josef Lange: An vielen Orten entstehen neue Studiengänge in Kooperation zwischen Universitäten und Fachhochschulen, aber auch künstlerischen Hochschulen, z.B. im Bereich Medien und Design. So weit ich das beurteilen kann, geht es in allen Fällen darum, dass an einem Ort oder benachbarten Orten vorhandene Potential zu bündeln für neue Studiengänge und zur Förderung des wissenschaftlichen Nachwuchses. Meistens reagieren die Hochschulen auf konkrete Anforderungen der beruflichen Praxis, der künftigen Arbeitgeber ihrer Absolventinnen und Absolventen. Hier zeigt sich sehr praktisch, wie notwendig die Forderung der Wissenschaftspolitik an die Hochschulen ist, bei der Konzeption neuer Studiengänge die Belange des Arbeitsmarktes nachhaltig zu berücksichtigen. Hochschulen bilden inzwischen 35 bis 40 Prozent eines Altersjahrgangs aus. Dieser Anteil geht weit über den Bedarf der Forschung hinaus. Hochschulstudium und -lehre vermitteln eine hoch qualifizierte Berufsausbildung. Dies zu tun ist eine der Erwartungen der Gesellschaft an öffentlich finanzierte Hochschulen.

HSW: Sind die jeweils geplanten Kooperationen und Vernetzungen überhaupt das richtige Instrument - bezogen auf die verfolgten Ziele?

Josef Lange: Bei der Orientierung am Arbeitsmarkt ist Kooperation vor Ort allemal das richtige Instrument. Niemandem wäre zu vermitteln, wenn bereits in einer Institution vorhandene Kapazität an demselben Ort in einer anderen Institution parallel erneut aufgebaut würde. Für eine solche Art von institutionellem Egoismus fehlt in Öffentlichkeit und Parlamenten jedes Verständnis. Die Verweigerung von Kooperation angesichts zu lösender Probleme und zu deckender Studienplatznachfrage wäre politisch und publizistisch nicht vermittelbar.

HSW: Sind die eingeschlagenen Wege und die Etappen in den Entwicklungsprozessen im Licht eines Change Management aus Ihrer Sicht richtig gewählt?

Josef Lange: Auch bei der Entwicklung von Studiengängen sind unterschiedliche Wege denkbar und werden unterschiedliche Wege gegangen – von Initiativen Einzelner in Fakultäten über Fakultätsinitiativen bis zu Anstößen zukunftsorientierter Hochschulleitungen und Anforderungen aus der beruflichen Praxis sowohl von Verbänden als auch großer Einzelunternehmen. Wenn Berührungspunkte zwischen Hochschulen und zwischen Hochschulen und Unternehmen abgebaut sind, fällt die Kooperation leichter. Aber auch hier hängt der Erfolg entscheidend von den handelnden Personen ab.

HSW: Sind - soweit absehbar - die richtigen Akteure beteiligt? Fehlt eine Seite?

Josef Lange: s.o.

HSW: Wie sehen Sie die Rolle des Staates in diesen Prozessen?

Josef Lange: In diesen Feldern hat staatliche Administration die Aufgabe, rechtliche Hindernisse der Vergangenheit auszuräumen, hat die Politik die Aufgabe zur Zusammenarbeit zu ermuntern. Die zu lösenden Probleme im globalen wissenschaftlichen und wirtschaftlichen Wettbewerb sind nicht eindimensional. Deshalb sind unterschiedliche Partner zur Lösung erforderlich. Notwendig ist die gegenseitige Achtung von Eigenständigkeit in der Differenz der Aufgaben und Institutionen.

HSW: Was wird staatlicherseits an Unterstützung gebraucht:

- politisch (auch symbolisch-verbale Unterstützung)?
- rechtlich (Änderung von Gesetzen, begleitende oder grundlegende Verordnungen)?
- finanziell?

Josef Lange: Rechtlich sind kooperative Studiengänge in Niedersachsen längst möglich. Sollte sich herausstellen, dass an der einen oder anderen Stelle Probleme bestehen, so ist es Aufgabe des Staates, sie so rasch und unkompliziert wie möglich zu beheben. Die symbolische Unterstützung von Veränderungsprozessen in Hochschulen durch die Politik wird in manchen Hochschulen gut aufgenommen oder gar eingefordert. In manchen Hochschulen scheint noch die Auffassung vorzuherrschen, der Staat solle nur das Geld zur Verfügung stellen und den Hochschulen alle

Freiräume auch in der Verausgabung öffentlicher Mittel einräumen und sich ansonsten aus Hochschulen und Hochschulpolitik heraushalten. Diese Position verkennt, dass die Finanzierung der Hochschulen aus öffentlichen Mitteln geschieht und dass die Hochschulen über deren Verwendung rechenschaftspflichtig sind. Autonomie der Hochschulen heißt vor allem Eigenverantwortung, die auch wahrgenommen werden muss. Es sollte ein Warnsignal für die Hoch-

schulen sein, wenn in Politik und Öffentlichkeit die Frage gestellt wird, ob die Hochschulen hinreichend autonomiefähig sind, weil sich zu viele ihrer öffentlichen Äußerungen auf die Forderung nach mehr Geld konzentrieren.

HSW: Herr Lange, wir danken Ihnen für diese Stellungnahmen.

René Krempkow

Leistungsbewertung, Leistungsanreize und die Qualität der Hochschullehre Konzepte, Kriterien und ihre Akzeptanz



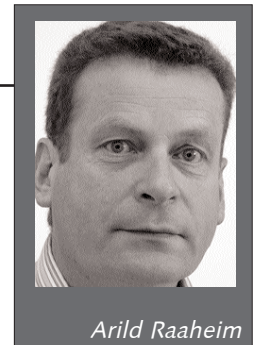
ISBN 3-937026-52-5,
Bielefeld 2007,
297 Seiten, 39.00 Euro

Mehr als eineinhalb Jahrzehnte sind vergangen, seit das Thema Bewertung der Hochschulleistungen und dabei vor allem der „Qualität der Lehre“ in Deutschland auf die Tagesordnung gebracht wurde. Inzwischen wird eine stärker leistungsorientierte Finanzierung von Hochschulen und Fachbereichen auch im Bereich der Lehre immer stärker forciert. Bislang nur selten systematisch untersucht wurde aber, welche (auch nicht intendierten) Effekte Kopplungsmechanismen zwischen Leistungsbewertungen und Leistungsanreizen wie die Vergabe finanzieller Mittel für die Qualität der Lehre haben können. Für die (Mit-)Gestaltung sich abzeichnender Veränderungsprozesse dürfte es von großem Interesse sein, die zugrundeliegenden Konzepte, Kriterien und ihre Akzeptanz auch empirisch genauer zu untersuchen. Nach der von KMK-Präsident Zöllner angeregten Exzellenzinitiative Lehre und der vom Wissenschaftsrat angeregten Lehrprofessur sowie angesichts des in den kommenden Jahren zu erwartenden Erstsemesteransturms könnte das Thema sogar unerwartet politisch aktuell werden.

Im Einzelnen werden in dieser Untersuchung die stark auf quantitative Indikatoren (v.a. Hochschulstatistiken) bezogenen Konzepte zur Leistungsbewertung und zentrale Konzepte zur Qualitätsentwicklung bezüglich ihrer Stärken und Schwächen sowie Weiterentwicklungsmöglichkeiten diskutiert. Bei der Diskussion von Leistungsanreizen wird sich über den Hochschulbereich hinaus mit konkreten Erfahrungen in Wirtschaft und öffentlicher Verwaltung auseinandergesetzt – auch aus arbeitswissenschaftlicher und gewerkschaftlicher Sicht. Bei der Diskussion und Entwicklung von Kriterien und Indikatoren zur Erfassung von Qualität kann auf langjährige Erfahrungen und neuere Anwendungsbeispiele aus Projekten zur Hochschulberichterstattung mittels Hochschulstatistiken sowie Befragungen von Studierenden und Absolventen sowie Professoren und Mitarbeitern zurückgegriffen werden. Abschließend werden Möglichkeiten zur Einbeziehung von Qualitätskriterien in Leistungsbewertungen und zur Erhöhung der Akzeptanz skizziert, die zumindest einige der zu erwartenden nicht intendierten Effekte und Fehlanreizwirkungen vermeiden und damit zur Qualität der Lehre beitragen könnten.

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

Arild Raaheim



Arild Raaheim

Eine Studie über die Interrater-Reliabilität zur Prüfung beim Erstjahresstudium der Psychologie*

Sieben Prüfer wurden gebeten, 50 Prüfungsarbeiten nacheinander zu beurteilen und zu zensieren. Die Arbeiten wurden von Studierenden im 1. Studienjahr Psychologie an der Universität von Bergen erstellt. Damit sollte die „Interrater reliability“ ermittelt werden. Eine neue Gruppe von 4 sehr erfahrenen Prüfern wurde berufen, diejenigen Arbeiten zu benoten, deren Ausgangsnoten am meisten divergierten. Die neuen Prüfer wurden gebeten, diejenige Note bei jeder Arbeit festzusetzen, von der sie meinten, dies sei die jeweils „zutreffende“ Note. Die Interrater-Korrelationen für die 7 ursprünglichen Prüfer lagen zwischen .55 und .91. Zugleich stellten die Forscher eine Differenz von 10 Zehntel oder mehr zwischen der besten und schlechtesten Note in 45% der Fälle fest. Die „korrekte“, von den erfahrenen Prüfern festgesetzte Note - so zeigte sich - variierte sehr stark.

1. Einleitung¹

Die Erstjahresstudie zum Thema Psychologie dient bei der norwegischen Universität als Aufnahmegrundlage für das Berufsstudium. Die zum jeweiligen Zeitpunkt mit den besten Noten ausgezeichneten Studenten, werden berücksichtigt, sofern die Anzahl der Plätze ausreicht. Die Studenten werden nach ihren Durchschnittsnoten mit zwei Dezimalstellen eingestuft. An der Universität von Bergen liegt die obere Grenze für die Aufnahme in den letzten Semestern bei etwa 2,10.² Es ist offensichtlich, dass eine solche Situation besondere Anforderungen sowohl an den Studenten als auch an die Verwaltungsmitarbeiter dieser Rangordnung stellt. Davon sind nicht zuletzt alle Beurteiler und Lektoren betroffen. Wenn Hundertstel entscheidend sind, werden Zehntel sehr wichtig. Man sollte daher annehmen, dass auf der Systemseite entscheidender Wert auf den Einsatz erfahrener, hochqualifizierter Lektoren gelegt wird. Außerdem sollte man davon ausgehen, dass im Fachbereichsumfeld eindeutige Vorgehensweisen für die Zensurarbeit vorhanden sind, u.a. bei der Sicherstellung einer Form zur Normierung verschiedener Antworten. Sofern keine absoluten Kriterien dafür vorliegen, was eine gute Antwort ist (im Unterschied zu einer sehr guten oder gar schlechten Antwort), wird eine solche Normierung in vielen Hinsichten entscheidend. Mögliche negative Effekte der idiosynkratischen Korrekturstrategien bei den Lektoren ließen sich damit verringern. Angesichts der hohen Anzahl von Prüfungsantworten in jedem Semester im Erstjahresstudium der Psychologie (etwa 2.000 an der Universität von Bergen) liegt es auf der Hand, dass praktische Aspekte andere Dinge in den Schattensetzen stellen, wenn es darum geht, welche Lektoren man einsetzt. In erster Linie geht es darum, viele zu fragen und alle

The lack of reliability in the rating of exams is a phenomenon that is known but relatively unworked. *Arild Raaheim*, Professor for Psychology at the University of Bergen (Norway) made **An investigation of interrater reliability at a first-year exam in psychology**. Therefore seven examiners were asked to mark fifty examination papers produced by first-year students in psychology at the University of Bergen, in order to investigate the interrater reliability. A new group of four very experienced examiners was appointed to mark the papers where the marks diverged the most. The new examiners were asked to present what they believed to be the „correct“ mark in each case. The interrater correlations for the seven examiners range between .55 and .91. At the same time there is a difference of ten 10ths or more between the best and the lowest mark in 45% of the cases. The „correct“ mark, as indicated by the experienced examiners, proved to vary very much.

zu nehmen, die „Ja“ sagen. Außerdem ist es eine Tatsache, dass keine eindeutigen Verfahren für die Korrekturarbeit vorliegen. Dem steht eine ganz klar formulierte Zielsetzung entgegen, dass eine Lektoratsanleitung nicht auszuarbeiten ist. Der einzelne Beurteiler/Lektor ist mit anderen Worten selbst und seinen eigenen Bewertungen überlassen. Etwa parallel zur Einteilung der Erstjahresstudie in kleinere Einheiten (Module) wurde die Anzahl der Kommissionsmitglieder von drei auf zwei reduziert. Diese Verringerung wurde in erster Linie aus wirtschaftlichen Gründen vorgenommen. Wer früher zwei achtstündige schriftliche Prüfungen und anschließend eine mündliche Prüfung abzulegen hatte, steht nach der Änderung vor vier schriftlichen vierstündigen Prüfungen. Gleichzeitig hat sich die Anzahl der Studenten in diesem Fach stark vermehrt. Heute sind etwa 1.000 Studenten im ersten Jahr des Psychologiestudiums an der Universität von Bergen eingeschrieben. Auf Anfrage wurde mitgeteilt, dass die psychologische Fakultät (im Jahr 1997) etwa drei Millionen norwegische Kronen für die Prüfungsdurchführung benötigt. Die (überraschend geringe) Kritik von Seiten der Studenten richtete sich hauptsächlich

* Dieser Artikel wurde bereits in der Zeitschrift der norwegischen Psychologengemeinschaft im Jahr 2000 unter der Nummer 37, 203-213 veröffentlicht. Die vorliegende Version wurde von der Zeitschrift genehmigt.

¹ Seit der Veröffentlichung dieser Studie in norwegischer Sprache haben die norwegischen Universitäten eine neue Notenskala eingeführt. Die neue Skala (ECTS-Skala) umfasst sechs Noten von A (Bestnote) bis F (nicht bestanden).

² Die ursprüngliche Notenskala erstreckte sich von 1,0 (Bestnote) bis 6,0, wobei die Noten im Bereich 1,0 bis 4,0 mit Zehnteln unterteilt wurden (1,1, 1,2, 1,3 usw.). Die Note 2,5 entspricht der Bewertung „cum laude“.

auf den Zeitaufwand bei Beschwerden. Die Sachbearbeitung nahm soviel Zeit in Anspruch, dass der Zug für die Studenten bereits abgefahren war, deren Beschwerde schließlich als berechtigt anerkannt und deren Notendurchschnitt somit um die erforderlichen Hundertstel angehoben wurde. Die Kritik, die in mehreren Fällen in der Presse veröffentlicht wurde, richtete sich gegen die empfundene Ungerechtigkeit, dass man keinen garantierten Platz bei der nächsten Aufnahme im nachfolgenden Semester erhielt. Es ist ein Widerspruch in sich, dass zentrale Fachpersonen (und Studenten) anscheinend das gegenwärtige Prüfungssystem akzeptieren, ohne Fragen nach den grundlegenden Voraussetzungen zu stellen, auf denen das System basiert, und ohne zu zeigen, dass ihnen ganz offensichtliche Schwächen nicht gleichgültig sind. Die Frage, welche Funktionen eine Prüfung hat bzw. haben sollte, habe ich an anderer Stelle behandelt. Dort wird auch die Problematik mit der Aufnahme zum Berufsstudium der Psychologie aufgegriffen und ein konkreter Vorschlag zu einer alternativen Aufnahmeform vorgelegt (Raaheim 1998). Hier liegt der Schwerpunkt bei Fragen, die sich damit befassen, wie man sich auf Beurteilungen der Prüfungsleistungen von Studenten verlassen kann. Wie zuverlässig (reliable) sind die vorgenommenen Beurteilungen? Die grundsätzliche Frage in Verbindung mit der Reliabilität in diesem Zusammenhang lautet wie folgt: Würden andere Beurteiler (Lektoren) sich meiner Beurteilung der Leistung von Studenten anschließen? Oder können wir die Frage etwas anders stellen und vielleicht etwas provokativer formulieren, wie das u.a. auch Rowntree (1987) tut: Wäre ich mit mir selbst einverstanden, wenn man mich bitten würde, dieselben Studenten und dieselben Aufgaben erneut zu beurteilen?

1.1 Interrater-Reliabilität

Die meisten Beurteiler/Lektoren haben sich vermutlich in dem einen oder anderen Zusammenhang diese Frage schon selbst gestellt, inwieweit Laune, Motivation, Tageszeit der Beurteilung, Handschrift, Reihenfolge der Antworten etc. sich auf die Beurteilung auswirken. Gleichzeitig scheint eine Form des Einverständnisses in akademischen Kreisen vorhanden zu sein, dass die Beurteilungen von Leistungen der Studenten „richtig“ sind, dass also qualifizierte Beurteiler eine Arbeit mehr oder weniger gleich betrachten und bewerten. Folglich liegt eine hohe Interrater-Reliabilität vor. Forschungsergebnisse weisen aus, dass dies absolut nicht immer zutrifft, sondern vielleicht eher sogar die Ausnahme ist (Fleming 1999; Rowntree 1987). Rowntree (1987) bringt dies wie folgt zum Ausdruck:

„Abundant evidence attests to the falsity of such assumptions ... Indeed, we might well substitute the verb *vident* for *faciunt* in the 'anti-behaviourist' Roman saying mentioned earlier, making it: *Si duo idem vident, non est idem* – if two people see the same thing, it is not the same thing“ (S. 191). Rowntree (1987) beschreibt verschiedene Studien, die über den Bereich Interrater-Reliabilität vorgenommen worden sind. Eine der ersten bekannten Untersuchungen wurde in Oxford im Jahr 1889 von F. Y. Edgeworth, Professor der Politikwirtschaft, durchgeführt. Er veröffentlichte „A piece of Latin prose“ im *Journal of Education* und forderte qualifizierte Beurteiler auf, diese so zu beurteilen, als ob es von einem Studenten geschrieben worden wäre. Insgesamt 28 „hochkompetente Prüfer“ folgten der Aufforderung und

beurteilten die Aufgabe mit insgesamt 15 verschiedenen Noten von 45 bis 100 (auf einer Skala mit 100 als Bestnote und einem Durchfallgrenzwert von normalerweise 45 oder 50). Es wurde auch auf eine Studie von Ben Wood verwiesen, die im Jahr 1921 durchgeführt wurde. Dabei ging es darum, tatsächliche Prüfungsantworten im Fach Geschichte zu bewerten. Einer der sechs an der Beurteilung beteiligten Professoren ging über das reine Lesen der Antworten der Studenten hinaus. Für sich selbst hatte er eine Version angefertigt, die er für die ideale Beantwortung einer solchen Aufgabe hielt. Durch ein Versehen gelangte diese „ideale Antwort“ unter die Antworten der Studenten. Ein anderer Beurteiler, der nichts davon wusste, dass diese Antwort von einem Kollegen stammte, ließ diese Beantwortung durchfallen. Im Rahmen des Standardkontrollverfahrens wurde auch die Note der „idealen Antwort“ einer doppelten Prüfung unterzogen. Das Ergebnis bestand aus verschiedenen Noten zwischen 45 und 90. Die möglicherweise am meisten diskutierte Studie dieser Art wurde von Diederich (siehe auch Rowntree 1987) vorgenommen. Hier sollten 53 Experten 300 Antworten (Kurzantworten) beurteilen, die von Erstjahresstudenten geschrieben worden waren. Die verwendete Notenskala war in neun Noten unterteilt. Die Ergebnisse wiesen aus, dass keine Antwort weniger als fünf verschiedene Noten erhalten hatte, und sogar 34% wurden mit allen neun Noten beurteilt. Abschließend soll noch eine Studie von Hartog und Rhodes aus der Mitte der 30-er Jahre erwähnt werden. Dabei wurde nicht nur eine ziemlich deutlich fehlende Übereinstimmung bei den Bewertungen der 15 verschiedenen Beurteiler festgestellt. Es konnte sogar dokumentiert werden, dass die Note davon abhing, wann die Antwort gelesen wurde. Die beurteilten Antworten sollten als „durchgefallen“, „bestanden“ oder „bestanden mit Auszeichnung“ eingestuft werden. Es stellte sich heraus, dass in mehr als 50% der Fälle dieselbe Antwort mit allen drei Beurteilungen bedacht wurde. Als die Beurteiler ein Jahr später gebeten wurden, dieselben Antworten erneut zu beurteilen (ohne dass dies vorher mitgeteilt wurde), stellte sich heraus, dass sie nicht einmal mehr mit sich selbst einig waren. In erstaunlich vielen Fällen (83) wurde die Bewertung von „durchgefallen“ in „bestanden“ und umgekehrt geändert. In neun Fällen erfolgte eine Änderung über zwei Kategorien hinweg (von „durchgefallen“ zu „bestanden mit Auszeichnung“ oder umgekehrt) (Rowntree 1987). Ein ähnliches Ergebnis wird von Gabb (1980) beschrieben. Dabei ergab sich eine extrem niedrige Korrelation von 0,28, wobei dieselben Lektoren dieselbe Art von Antworten zu zwei verschiedenen Zeitpunkten beurteilt hatten. In Norwegen sind mehr Untersuchungen bekannt, die sich ausführlicher mit der Übereinstimmung von Bewertungen verschiedener Beurteiler im Erstjahresstudium der Psychologie befassen (Engvik/Kvale/Havik 1970; Havik/Teigen/Kvale 1969; Lundal/Manger/Mæland 1991; Teigen 1986; Teigen/Kvale/Tschudi 1970). Havik et al. (1969) nahmen eine Bewertung der Intra- und Interrater-Reliabilität im Rahmen verschiedener Prüfungsausschüsse an der Universität von Oslo vor. Unter Intra-Reliabilität wird die Übereinstimmung bei Antworten am 1. und 2. Tag verstanden, während Interrater-Reliabilität sich auf die Übereinstimmung der drei Ausschussmitglieder untereinander bezieht. Die Ergebnisse wiesen aus, dass sowohl die Intra- als auch die Interrater-Reliabilität ausgesprochen va-

riabel und durchgehend niedrig war (Intra-Reliabilität nur 0,23 und Interrater-Reliabilität nur 0,33). Teigen et al. (1970) nahmen eine umfangreiche Untersuchung der Interrater-Reliabilität vor, bei der verschiedene Ausschüsse dieselben Prüfungsaufgaben beurteilten. Neben den Ausschüssen aus Fachleuten mit Anschluss an die Universitäten von Oslo und Bergen wurden hier auch andere Kommissionen eingesetzt, die sich aus Studenten (einige Anfänger, einige gerade fertig mit dem ersten Studienjahr) der Universitäten von Oslo, Bergen und Trondheim zusammensetzten. Diese Studie (Teil 1) ergab eine erheblich größere Übereinstimmung bei den Bewertungen als dies bei der Untersuchung von Havik et al. (1969) der Fall gewesen war. Es ergab sich eine ausgesprochen hohe Übereinstimmung bei der endgültigen Benotung mit Korrelationen im Bereich von 0,825 und 0,925. Die Kontrollausschüsse (hauptsächlich aus Studenten bestehend) stellten eine Tendenz zur besseren Benotung und zur etwas geringeren Ausnutzung der Notenskala fest. In keinem einzigen Fall erhielt ein Kandidat mit „cum laude“ in einer Kommission eine schlechtere Benotung als 2,8 in einer anderen. Ansonsten wurde hier festgestellt, dass alle Korrelationen zwischen Einzelbeurteilern niedriger ausfallen als Korrelationen zwischen den Ausschussbenotungen (im Bereich 0,514 bis 0,748, wobei verschiedene Lektoren dieselben Aufgaben benoten). In einer neuen Untersuchung greift Teigen (1986) die Benotung aller Prüfungsausschüsse für das Erstjahrestudium der Psychologie an den Universitäten von Bergen und Oslo im Herbst 1985 auf (fünf Kommissionen in Bergen und fünf in Oslo). Erneut zeigt sich, dass die Übereinstimmung der Benotungen durchgehend hoch ist (dieselbe Aufgabe wird von unterschiedlichen Lektoren beurteilt). In Bergen liegen 80% aller Einzelkorrelationen über 0,70. In Oslo ist das Ergebnis nicht ganz so positiv, dort liegen 40% der Einzelkorrelationen über 0,70. Die Intra-Reliabilität war mit durchschnittlich 0,66 (Bergen) und 0,63 (Oslo) wiederum etwas niedriger. Diese Studie belegt, dass eine Abweichung von mehr als fünf Zehnteln eher zu den Seltenheiten gehört. In 60 bis 66% der Fälle beschränkt sich die Abweichung auf bis zu zwei Zehntel. Die Untersuchung von Lundal et al. (1991) wurde durchgeführt, nachdem die Anzahl der Mitglieder in der Prüfungskommission von drei auf zwei reduziert wurde, aber vor der Modularisierung des Fachgebiets. Die Ergebnisse dieser Untersuchung ergeben ein etwas anderes Bild als das der Studie von Teigen et al. (1970) und Teigen (1986). Hier wird auch eine große Übereinstimmung zwischen den beiden Kommissionsmitgliedern deutlich. Die Uneinigkeit ist am ersten Tag am größten. Bei 25% der Antworten betragen die Unterschiede in der Bewertung mehr als vier Zehntel. Bei 10% der Fälle liegen die Unterschiede bei mehr als sechs Zehnteln, wobei die größte Abweichung 24 Zehntel beträgt. Die Ergebnisse fallen bei den Zensuren am zweiten Tag etwas besser aus. Der größte Unterschied liegt hier bei 22 Zehntel, wobei der Durchschnitt drei Zehntel beträgt. Übereinstimmend mit den Erkenntnissen von Teigen (1986) zeigt auch diese Untersuchung, dass die Fähigkeit der Studenten, ihre Leistungen im Rahmen einer Prüfung zu beurteilen, nicht besonders gut ist. Der Übereinstimmung zwischen den eigenen Aussagen zu den realistischen Erwartungen und der endgültigen Benotung liegt bei 0,15. Auch zwischen den Ergebnissen einer Testprüfung und der Prüfungsbenotung besteht nur ein relativ geringer Zusammenhang (0,39).

1.2 Auswirkung der idiosynkratischen Korrekturstrategien

Wenn wir als Fachleute eine Benotung auf uns nehmen oder auferlegt bekommen, wollen wir das Lektorat selbst auf unterschiedliche Weise gestalten. So wollen wir z.B. die Prüfungsantworten nach Geschmack organisieren und die dünnsten/dicksten zuerst lesen. Andere organisieren die Antworten gern nach Aufgabenbereichen und lesen alle Aufgaben vom Typ X, bevor sie die Aufgaben vom Typ Y in Angriff nehmen. Wiederum andere lesen zunächst die Antworten mit schöner und leicht zu lesender Handschrift. Viele machen es wie ich und lesen die Antworten in der Reihenfolge des Stapels. Wir sehen hier, dass jeder seine Bewertung auf der Grundlage eigener Arbeitsmethode festlegt und dabei bestimmte Voraussetzungen berücksichtigt, die sich auf die Benotung auswirken können. Wenn zwei Personen die Antworten nach unterschiedlichen „Rezepten“ lesen, ist schon ein Unterschied bei der Benotung vorhanden, der in den Noten zum Ausdruck kommen kann.

Die Frage nach möglichen Fehlerquellen in Verbindung mit den Bewertungen der Prüfungsantworten der Studenten wird u. a. von Fleming (1999) sowie von McDowell und Sambell (1999) behandelt. Fleming (1999) legt eine Übersicht über die hier in Frage kommenden Fehlerquellen vor. Vermutungen zum Geschlecht der Studenten auf Seiten des Lektors hat in einzelnen Fällen (aber nicht immer) bei der Beurteilung des Verständnisses der Studenten eine Rolle gespielt. In Einzelfällen konnte festgestellt werden, dass bei der Annahme, dass die Arbeit von einem Mann geschrieben wurde, zu positiveren Benotungen gegriffen wurde. Fleming verweist auch auf Studien, die belegen, dass der Lektor (sowohl Frauen als auch Männer) eine Tendenz hat, fälschlicherweise anzunehmen, dass eine undeutliche Handschrift von Männern stammt. Es ist auch dokumentiert worden, dass Personen, denen die Studenten bekannt sind, eine Tendenz aufweisen, bessere Noten zu verteilen als Leute, die diese nicht kennen. In Flemings Zusammenfassung wird auch Bezug auf Untersuchungen genommen, die belegen, dass die Frage nach dem Wie, also nach der Reihenfolge beim Lesen der Antworten, nicht ohne Bedeutung ist. So steht u. a. fest, dass eine Antwort mit mittlerer Qualität („average essay“) im Anschluss an eine Reihe von guten Antworten schlechtere Noten erhält als wenn sie nach einer Reihe von schlechten Antworten gelesen wird. Studien von Sprouse und Webb (siehe auch Fleming 1999) belegen, dass Antworten mit deutlicher Handschrift bessere Noten erhalten als Antworten mit schlechter, undeutlicher Handschrift.

Im Zeitraum seit der Durchführung der Studien von Teigen (1986) und Lundal et al. (1991) haben sich wie bereits erwähnt große Veränderungen im Erstjahrestudium der Psychologie vollzogen. Die Anzahl der Studenten ist dramatisch angestiegen. Das Fachgebiet wurde modularisiert und wird in vier fünfwöchigen Kursen mit einer jeweils vierstündigen, schriftlichen Abschlussprüfung durchgeführt. Die Besetzung der Prüfungskommission mit zwei Mitgliedern ist als dauerhafte Einrichtung etabliert. Gleichzeitig gehört eine Ordnung der Vergangenheit an, bei der alle mit der Note „cum laude“ garantiert einen Platz für das Psychologiestudium erhielten (sofern sie bereit waren, zu warten). Heute wird ein System praktiziert, bei dem die Studenten um Hundertstel konkurrieren. Die Frage nach der Übereinstimmung der Benotungen

der heutigen Lektoren und die Frage nach eventuell vorhandenen (großen) Unterschieden bei der Benotung verschiedener Kommissionen werden auf dieser Grundlage erneut aufgeworfen. Eigene Erfahrungen aus nahezu 20 Jahren Arbeit mit Beurteilungen in diesem Fachbereich legen die Vermutung nahe, dass nach den genannten Umstrukturierungen eine Veränderung in negativer Richtung stattgefunden hat. Dies wurde besonders im letzten Semester deutlich. Hier war der Abstand zwischen meinen eigenen Noten und denen der Lektoren in mehreren Fällen beunruhigend groß (bis zu einer ganzen Note bei derselben Aufgabe). Daher wurde die Entscheidung getroffen, eine Studie durchzuführen, deren Zielsetzung darin besteht, die Interrater-Reliabilität beim Erstjahrestudium der Psychologie zu untersuchen.

2. Methode

Ein Satz von 50 Prüfungsantworten aus dem Erstjahrestudium der Psychologie (PS101, Herbst 1998) an der Universität von Bergen wurde im Frühjahr 1999 fünf unabhängigen Beurteilern zugeschickt. Dies geschah, nachdem ein normales Lektorat stattgefunden hatte. Getrennte Noten der Mitglieder in der ursprünglichen Prüfungskommission sowie die endgültigen Noten wurden ebenfalls erfasst. Wir haben somit sieben Personen, die sich unabhängig von einander mit den Antworten beschäftigt haben. Alle Personen stehen auf der Liste offiziell zugelassener Lektoren der psychologischen Fakultät und werden auch bei der normalen Benotung in diesem Fachbereich eingesetzt. Vier der Lektoren sind sehr erfahren und verfügen über eine durchschnittliche Erfahrung von nahezu 20 Jahren. Die anderen drei verfügen über weniger Erfahrung (zwischen einem und zehn Jahren). Wir haben diese (5) Personen angesprochen, damit sie (zusammen mit den ursprünglichen zwei) eine Art „Querschnitt“ des eingesetzten Lektoratteams sicherstellen sollten. Die fünf Beurteiler, denen die Antworten nach der Vergabe der normalen Note zugestellt wurden, erhielten die Mitteilung, dass die „Zensur“ zu einer Art „Kontrollstudie“ im Rahmen eines Testprojektes in Verbindung mit PS104 der Universität von Bergen im Frühjahr 1999 gehörte. Dabei erhielten die Studenten Gelegenheit, die traditionelle „Schulprüfung“ durch eine Struktur zu ersetzen, bei der sie stattdessen zwei schriftliche Arbeiten im Lauf des Semesters abgeben konnten. Statt der Notenzahlen wurde eine fünfteilige Skala verwendet („nicht bestanden“, „bestanden ausreichend“, „bestanden befriedigend“, „bestanden gut“ und „bestanden sehr gut“). Insgesamt 53 Studenten nahmen an dem Testprojekt teil (siehe Raaheim

1999). Unsere Lektoren erhielten die Angabe, dass die von ihnen zu beurteilenden 50 Arbeiten ursprünglich von den Studenten im Rahmen des Testprojektes geschrieben worden waren, als sie die Prüfung von PS101 ablegten. Die Arbeiten waren vorher bereits von verschiedenen Kommissionen benotet worden. Sie waren jetzt mit der Absicht auf eine Kommission konzentriert worden, eventuelle Kommissionsauswirkungen auf die Benotung zu überprüfen. Diese Erklärung wurde von allen Lektoren akzeptiert. Sie hatten genauso lange Zeit wie bei der normalen Beurteilungsarbeit, und alle erhielten eine Vergütung nach den gültigen Sätzen. Zusammen mit den Arbeiten erhielten die Lektoren eine Liste der Kandidaten, die sie jeweils direkt an mich gesendet haben. Nachdem alle Lektoren ihre Notenlisten abgegeben hatten, wurden sieben Arbeiten herausgesucht und an vier der erfahrensten Prüfer von Norwegen geschickt. Dabei handelte es sich um Professoren, die 30 bis

Tabelle 1: Notenverteilung von 7 Lektoren sowie Prüfungsnote

Kandidat	Lektor 1	Lektor 2	Lektor 3	Lektor 4	Lektor 5	Lektor 6	Lektor 7	Prüfungsnote
1	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
2	n. b.	n. b.	n. b.	n. b.	n. b.	3,6	2,9	n. b.
3	2,5	3,2	3,5	2,6	2,7	3,6	2,8	3,0
4	2,8	2,5	2,9	2,8	3,3	3,2	3,9	2,8
5	3,2	2,8	n. b.	3,5	3,5	3,6	3,2	n. b.
6	2,9	2,6	3,1	2,9	2,6	2,9	2,7	3,0
7	2,8	2,5	3,3	2,8	3,3	2,6	2,6	3,0
8	2,8	2,5	2,8	2,4	3,0	2,6	2,8	2,8
9	2,7	2,5	2,7	2,4	2,6	2,7	2,8	2,8
10	2,8	2,2	2,8	2,7	2,7	2,6	3,0	2,8
11	3,0	2,8	n. b.	3,0	2,8	3,0	3,0	3,5
12	2,6	2,6	2,9	2,6	2,8	2,8	2,7	2,7
13	2,8	2,4	2,8	2,6	2,8	2,7	2,8	2,8
14	3,2	3,1	3,7	3,1	3,4	2,8	2,9	3,5
15	3,0	2,9	n. b.	2,7	3,3	2,7	2,8	3,5
16	2,6	2,6	3,2	2,6	3,0	2,6	2,7	2,9
17	3,3	3,2	n. b.	3,0	3,2	2,9	3,3	3,5
18	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	3,8	n. b.
19	3,0	2,5	n. b.	2,8	3,6	2,9	2,9	n. b.
20	3,0	n. b.	n. b.	2,9	3,7	3,4	3,2	3,8
21	2,9	2,8	n. b.	3,3	3,6	2,9	2,8	3,3
22	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
23	3,1	3,5	3,5	2,8	3,3	2,9	3,0	3,3
24	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	3,5	n. b.
25	n. b.	3,9	n. b.	3,8	n. b.	n. b.	n. b.	n. b.
26	2,8	2,6	n. b.	2,7	2,9	2,4	2,4	3,5
27	2,3	1,9	2,2	1,8	2,1	2,0	2,1	2,2
28	2,4	2,0	2,7	1,8	2,7	2,3	2,5	2,5
29	2,1	2,0	2,6	1,6	2,5	2,3	2,4	2,3
30	1,9	2,2	2,3	1,9	2,3	2,1	2,3	2,1
31	n. b.	2,9	n. b.	3,2	n. b.	n. b.	2,7	n. b.
32	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
33	2,6	2,5	2,9	2,4	2,5	2,8	2,5	2,7
34	n. b.	n. b.	n. b.	3,3	n. b.	3,7	3,9	n. b.
35	3,0	n. b.	3,6	2,9	n. b.	3,0	3,2	3,3
36	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
37	n. b.	3,4	n. b.	2,8	n. b.	3,6	2,4	n. b.
38	2,5	2,6	2,9	2,3	2,9	2,4	2,4	2,7
39	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
40	n. b.	3,5	n. b.	3,0	n. b.	n. b.	2,9	n. b.
41	2,9	2,6	3,0	2,8	2,5	2,8	2,5	2,9
42	2,9	2,6	n. b.	3,0	n. b.	3,3	2,9	3,5
43	2,8	3,8	3,3	3,5	n. b.	2,8	2,8	3,0
44	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
45	3,2	3,3	3,4	3,2	n. b.	3,1	3,3	3,3
46	2,4	2,2	2,5	1,7	2,3	2,3	2,0	2,4
47	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
48	2,6	1,8	2,9	2,0	2,9	2,9	2,0	2,7
49	3,5	3,7	n. b.	n. b.	n. b.	3,5	3,5	n. b.
50	2,0	2,2	2,7	1,8	2,1	2,2	2,2	2,3

Tabelle 2: Korrelation zwischen verschiedenen Prüfern sowie zwischen Prüfer und tatsächlichem Prüfungsergebnis

	Lektor 1	Lektor 2	Lektor 3	Lektor 4	Lektor 5	Lektor 6	Lektor 7	Prüfung
Lektor 1	1.00	.778	.739	.801	.832	.910	.702	.905
Lektor 2	.778	1.00	.649	.827	.793	.747	.742	.733
Lektor 3	.739	.649	1.00	.705	.721	.700	.555	.880
Lektor 4	.801	.827	.705	1.00	.783	.805	.789	.812
Lektor 5	.832	.793	.721	.783	1.00	.788	.650	.822
Lektor 6	.910	.747	.700	.805	.789	1.00	.774	.852
Lektor 7	.702	.742	.555	.789	.650	.774	1.00	.665
Prüfung	.905	.733	.880	.812	.822	.852	.665	1.00

wählt, bei denen die Beurteilungen der sieben Lektoren die größten Unterschiede aufwiesen. Die „Superprüfer“ wurden über den Zweck der Untersuchung informiert. Sie erfuhren außerdem, welche sieben Noten die Arbeiten erhalten hatten. Sie erhielten die Anweisung, die „richtige“ Note für die Arbeiten zu vergeben.

3. Ergebnisse

Tabelle 3: Fiktive Kommissionsnoten und tatsächliche Prüfungsnoten

Kandidat	Prüfer 1 und 2	Prüfer 6 und 7	Prüfer 4 und 6	Prüfer 3 und 4	Prüfer 2 und 6	Prüfer 2 und 7	Prüfer 1 und 3 (Prüfungsnote)
1	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
2	n. b.	3.2	n. b.	n. b.	n. b.	3.3	n. b.
3	2.7	3.1	3.0	3.0	3.4	3.0	3.0
4	2.6	3.5	3.0	2.8	2.8	3.0	2.8
5	3.0	3.4	3.5	n. b.	3.1	3.0	n. b.
6	2.7	2.8	2.9	3.0	2.7	2.6	3.0
7	2.7	2.6	2.7	3.0	2.5	2.5	3.0
8	2.7	2.7	2.5	2.6	2.5	2.6	2.8
9	2.6	2.7	2.5	2.5	2.6	2.6	2.8
10	2.5	2.8	2.6	2.7	2.4	2.5	2.8
11	2.9	3.0	3.0	3.5	2.9	2.9	3.5
12	2.6	2.7	2.7	2.7	2.6	2.6	2.7
13	2.6	2.7	2.6	2.7	2.5	2.6	2.8
14	3.1	2.8	2.9	3.4	2.9	3.0	3.5
15	2.9	2.7	2.7	3.5	2.8	2.8	3.5
16	2.6	2.6	2.6	2.9	2.6	2.6	2.9
17	3.3	3.1	2.9	3.5	3.0	3.2	3.5
18	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
19	2.7	2.9	2.8	3.5	2.7	2.7	n. b.
20	3.5	3.3	3.1	3.5	n. b.	3.7	3.8
21	2.8	2.8	3.1	n. b.	2.8	2.8	3.3
22	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
23	3.2	2.9	2.8	3.2	3.3	3.2	3.3
24	n. b.	4.0	n. b.	n. b.	n. b.	n. b.	n. b.
25	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
26	2.7	2.4	2.5	3.5	2.5	2.5	3.5
27	2.1	2.0	1.9	2.0	1.9	2.0	2.2
28	2.2	2.4	2.1	2.2	2.2	2.2	2.5
29	2.0	2.3	1.9	2.0	2.2	2.2	2.3
30	2.0	2.2	2.0	2.1	2.1	2.2	2.1
31	3.3	3.3	3.6	n. b.	2.8	2.8	n. b.
32	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
33	2.5	2.6	2.5	2.6	2.6	2.5	2.7
34	n. b.	3.8	3.5	n. b.	n. b.	n. b.	n. b.
35	3.5	3.1	2.9	3.2	3.5	3.7	3.3
36	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
37	3.5/n. b.	2.8	3.1	3.5	3.5	2.8	n. b.
38	2.5	2.4	2.3	2.6	2.5	2.5	2.7
39	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
40	n. b.	3.5	3.5	3.6	n. b.	3.2	n. b.
41	2.7	2.6	2.8	2.9	2.7	2.5	2.9
42	2.7	3.1	3.1	3.6	2.9	2.7	3.5
43	3.3	2.8	3.1	3.4	3.3	3.2	3.0
44	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
45	3.2	3.2	3.1	3.3	3.2	3.3	3.3
46	2.3	2.1	2.0	2.1	2.2	2.1	2.4
47	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.	n. b.
48	2.4	2.4	2.4	2.4	2.3	1.9	2.7
49	3.6	3.5	n. b.	n. b.	3.6	3.6	n. b.
50	2.1	2.2	2.0	2.2	2.2	2.2	2.3

Aus Tabelle 1 geht die Notenverteilung aller sieben Lektoren sowie die tatsächliche Benotung der einzelnen Kandidaten hervor (Prüfungsnote). Wie man in dieser Tabelle deutlich erkennen kann, variieren die Noten bei den Einzelantworten sehr stark. Drei Kandidaten fallen bei sämtlichen Lektoren aus formellen Gründen durch, weil sie nur auf eine oder zwei (von insgesamt drei) Fragen geantwortet haben. In der weiteren Darstellung werden diese drei Kandidaten nicht berücksichtigt. Also bleiben 47 Kandidaten übrig. Bei genauerer Betrachtung kann man feststellen, dass der Abstand zwischen bester und schlechtester Note bei 24 Kandidaten mehr als eine ganze Note beträgt. In fünf Fällen ist das ein Unterschied zwischen der Note 3 und „nicht bestanden“. Da man als Lektor die Noten von 4.0 bis durchgefallen nicht abstuft, ist es schwierig zu sagen, dass wir vor einer Kluft von mehr als zehn Zehnteln stehen, bei der die beste Note z.B. 3,2 und die schlechteste „nicht bestanden“ lautet. Wenn wir diese fünf Kandidaten ausklammern, sind noch 19 Kandidaten übrig, bei denen der Unterschied zwischen bester und schlechtester Note mehr als eine ganze Note beträgt, also 45,2% der Fälle (19 von 42). Wie aus Tabelle 1 ersichtlich wird, unterscheiden sich die Noten in einigen Fällen von 2,4 bis durchgefallen (Kandidaten 26 und 37). An anderer Stelle erstreckt sich der Abstand von 1,6 bis 2,6 (Kandidat 29). An wiederum anderer Stelle erstreckt sich der Abstand von 1,8 bis 2,9 (Kandidat 48). Bei drei Kandidaten (19, 26 und 37) schwanken die Noten zwischen gut und

40 Jahre an der Universität tätig gewesen sind und allesamt über umfangreiche Erfahrungen im Bereich Unterricht und Benotung im Erstjahresstudium der Psychologie verfügen. Diese „Superprüfer“ stammen jeweils aus einer der vier norwegischen Universitäten. Wir hatten Arbeiten ausge-

durchgefallen. Wie aus der Tabelle abzulesen ist, ist ein bestimmter Lektor (Nr. 3) ständig etwas strenger als die anderen. Dieser verwendet die Skala auch etwas weniger als andere.

Auch wenn der Abstand zwischen bester und schlechtester Note groß ist, erkennt man in Tabelle 2, dass die Interrater-Korrelation relativ hoch ist. Die niedrigste Korrelation liegt bei 0,55 (zwischen Lektor 7 und 3), während die höchste Korrelation 0,91 beträgt (zwischen Lektor 1 und 6). Alle unsere Lektoren lasen die verschiedenen Arbeiten separat und nahmen die Benotung vor, ohne mit einem anderen Lektor über die Note gesprochen zu haben. Ausnahmen sind die Lektoren 1 und 3, die in der ursprünglichen Prüfungskommission saßen. Was hier als Prüfungsnote ausgewiesen ist, steht mit anderen Worten für die gemeinsame Benotung der Kommission. Wenn man jetzt davon ausgeht, dass die Arbeiten der Kandidaten von zwei anderen Prüfern beurteilt worden wären, wie wäre dann die Note ausgefallen? Tabelle 3 enthält die Noten der 50 Kandidaten, wie sie hätten aussehen können, wenn zwei andere Lektoren die Beurteilung vorgenommen hätten. Die Kommissionsnote geht hier als Durchschnitt der beiden Lektoren auf Abrundung auf das beste Zehntel. Auch wenn es sich dabei um ein ausgesprochen übliches Verfahren handelt, leidet unser Beispiel natürlich auch unter der Schwäche, dass die beiden Lektoren einander nicht getroffen haben und unsere künstliche Kommissionsnote daher nicht unbedingt vollkommen korrekt ausfällt. Dennoch wagen wir die Behauptung, dass die Noten nicht wesentlich von unseren konstruierten Noten abgewichen wären, wenn die Lektoren sich tatsächlich abgesprochen hätten.

Darin sind mehrere interessante Aspekte zu beachten. So erkennt man z.B., dass die Note „Gut“ eindeutig am wenigsten in der tatsächlichen Prüfungskommission vorkommt. Sechs Kandidaten erhielten diese Auszeichnung. In unseren fiktiven Kommissionen variiert dieser Wert zwischen acht (Lektoren 3 und 4) und 13 (Lektoren 2 und 7). Gleichzeitig kommt die Note „Nicht bestanden“ eindeutig am häufigsten in der tatsächlichen Prüfungskommission vor. Am wenigsten Durchgefallene gibt es bei den Lektoren 6 und 7 (9). Ansonsten schwankt die Anzahl der Durchgefallenen zwischen 9 und 16. Wenn man sich einzelne Noten genauer ansieht, entdeckt man teilweise große Unterschiede zwischen der Prüfungskommission und der einen oder anderen Kommission. Kandidat 19 fiel bei der Prüfungskommission durch. Hier erkennen wir, dass die anderen Kommissionen zu ganz anderen Beurteilungen kommen (etwa 2,7). Ausnahmen sind hier die Lektoren 3 und 4 mit der Note 3,5. Kandidat 26 erhielt in der Prüfung die Note 3,5. Dieselbe Note bekommt er von der Kommission, die sich aus den Lektoren 3 und 4 zusammensetzt. In vier Fällen lautet das Ergebnis ansonsten „Gut“. Kandidat 48 erhielt in der Prü-

fung die Note 2,7. Wenn seine Arbeit von anderen Prüfern gelesen worden wäre, hätte der Kandidat höchstwahrscheinlich ein weitaus besseres Ergebnis erzielt. In unserem Beispiel erkennen wir, dass die Note 1,9 vergeben worden wäre, wenn die Lektoren 2 und 7 in der Kommission gesessen hätten. Ansonsten kann man einen Umstand herausarbeiten. Man erkennt, dass die tatsächliche Prüfungskommission vielleicht besonders streng gewesen ist. In den meisten Fällen hätten die Kandidaten eine bessere Note erhalten, wenn die Arbeit von anderen Lektoren beurteilt worden wäre. Das gilt auch für diejenigen mit den besten Noten der Prüfungskommission. In einigen Fällen ist der Unterschied sehr groß.

Aus Tabelle 4 geht die Notenverteilung der vier „Superprüfer“ hervor. Hier wird die Notenverteilung der sieben Arbeiten dargestellt, bei denen die Spanne zwischen den „normalen“ Lektoren am größten ausgefallen ist (Kandidaten 4, 19, 26, 31, 35, 37 und 48). In der Tabelle ist auch angegeben, wie der Kandidat in der Prüfung tatsächlich benotet wurde. Auch hier ist das Spektrum sehr breit gefächert. In drei von sieben Fällen beträgt der Unterschied mehr als eine ganze Note. Am größten ist die Abweichung bei Kandidat 35 mit der besten Note 2,7 und der schlechtesten „Nicht bestanden“. Alle „Superprüfer“ hatten unaufgefordert persönliche Kommentare beigefügt. In zwei Fällen sind diese ausgesprochen ausführlich ausgefallen. Wir halten diese Informationen für wichtig in dieser Diskussion über die Gründe für die großen Notenunterschiede und nehmen daher einige Zitate auf. Drei der Lektoren geben an, dass sie die Arbeiten blind gelesen haben (also ohne zu überprüfen, welche Note die ersten sieben verteilt hatten). Der vierte sagt zur eigentlichen Korrekturarbeit: „In meiner Benotung habe ich meiner Meinung nach die von anderen vorgeschlagenen Noten nicht berücksichtigt ...“

„Ich glaube, dass ein Grund für die Notenabweichungen darin zu suchen ist, dass die Aufgaben sehr breit gestreut sind und der Kandidat daher viele sehr allgemeine Dinge sagt, die den Lektor nicht besonders beeindruckt haben. Mit anderen Worten werden die Kandidaten nicht für ihr Wissen belohnt, zu jeder einzelnen Aufgabe (abgesehen von der über die Nervenzellen) hätten sie mehrere Stunden lang schreiben können.“ (Lektor 3).

„Ich kann verstehen, dass es häufig zu Abweichungen bei der Benotung kommen kann, z.B. wenn Zweifel darüber bestehen, ob eine ansonsten akzeptable Antwort der Aufgabenstellung streng genommen gerecht wird oder nicht. Das trifft selbstverständlich häufig zu, wenn jemand durchfällt. Kandidat ... (48) ... hat eine hervorragende Antwort auf die Neurologiefrage gegeben und auch auf die beiden „Psychologiethemata“ keine schwachen Antworten geliefert (2,5 bis 2,8). Die Berechnung der Gesamtnote hängt natürlich davon ab, welche zentrale Bedeutung man der Neuropsychologie im Studium beimisst. Wenn ein Kandidat aber 1/3 seiner Leistung mit einer solchen Frage abdecken kann, bedeutet das wohl auch, dass man diese Frage entsprechend vage beantwortet.“ (Lektor 4).

„Meine Benotungsvorschläge sind der beiliegenden Liste zu entnehmen. Ich muss jedoch hinzufügen, dass die den Kandidaten gestellten Fragen Spielraum

Tabelle 4: Notenverteilung bei den „Superprüfern“ sowie Prüfungsnote ausgewählter Kandidaten

Kandidat	Lektor 1	Lektor 2	Lektor 3	Lektor 4	Prüfungsnote
4	3,2	3,7	2,8	2,7	2,8
19	3,5	2,9	3,1	2,8	n. b.
26	2,8	2,6	2,5	2,9	3,5
31	2,9	3,5	n. b.	n. b.	n. b.
35	3,0	3,3	2,7	n. b.	3,3
37	n. b.	3,2	n. b.	n. b.	n. b.
48	2,3	2,8	2,2	2,2	2,7

für verschiedene Auslegungen bieten, was die Erwartungen auf die Themenbehandlung betrifft (neben einer kurzen Erklärung eines vorgegebenen Begriffs – in dieser Aufgabe das Wort Zuordnung) oder wie breit der Kandidat seine Antwort auslegen darf (wie in der Aufgabe über die Modifizierung des Verhaltens, die sich nicht ohne Weiteres auf eine „Verhaltensveränderung“ als technischer/fachlicher Begriff begrenzen lässt). Dadurch ergibt sich beim Beurteiler eine Meinung jenseits jeglicher Angemessenheit.“

„Kandidat ... (48) ... wurde offenbar von einem Lektor für seine durchgehend guten Kenntnisse belohnt. Wenn man keine Abzüge für die ausgesprochen weitläufige Antwort zum Thema Lehren vornimmt, lautet das Ergebnis „Gut“. Ich habe die Note 2,8 vergeben, weil ich „glaube“, dass die Aufgabensteller eine enger gefasste Antwort erwartet haben. Ob das aber richtig ist?“ (Lektor 2).

„Die Benotung war schon etwas schwierig. Sie gehörte nicht zu den schwierigsten meiner Laufbahn, aber ich verstehe, dass die Streuung bei Antworten wie diesen groß ausfallen kann. Einzelne Kandidaten beherrschen einiges vom Unterrichtsstoff, haben aber nicht auf die Fragestellung der Aufgabe geantwortet. Andere wiederum haben ihre Stärken bei den gegebenen Beispielen, haben offenbar viel verstanden, aber gleichzeitig den Unterrichtsstoff schlecht aufgenommen.“ (Lektor 1).

4. Diskussion

Es gibt mehrere Aspekte, die in einer Diskussion über die Ergebnisse dieser Untersuchung hervorgehoben werden müssen. Wenn verschiedene Lektoren dieselbe Arbeit (sehr) unterschiedlich bewerten, ist das nicht unbedingt (oder insgesamt) ein Anzeichen dafür, dass bei einem oder mehreren der eingesetzten Lektoren Inkompetenz vorliegt. Eindeutig zu erkennen ist jedoch, dass verschiedene Lektoren Wert auf unterschiedliche Dinge legen, wenn sie ihre Benotung vornehmen. Und so kann sich das entwickeln, wenn von Seiten des Systems keine klaren Kriterien vorliegen, was eine bestimmte Antwort enthalten muss (Prüferanweisung). Derartige Kriterien sollten die absolute Mindestanforderung sein. Damit ließe sich verhindern, dass die Lektoren die Lösung der Aufgabenstellung (verschieden) interpretieren müssten. Aber es ist selbstverständlich nicht besonders sinnvoll, mehr oder weniger detaillierte Prüferanweisungen auszuarbeiten, weil die Lektoren ohnehin ihrer eigenen Einschätzung folgen und mit einer solchen Anweisung nur eine geringfügige Verpflichtung verbinden. Eine ausführliche Prüferanweisung kann allerdings auch verhindern, dass eine andersartige, aber richtige Antwort akzeptiert oder gut benotet wird. Mit anderen Worten bleibt die Frage unbeantwortet, ob es immer ein Vorteil ist, eine ausführliche Prüferanweisung anzubieten. Die Benotung kann allerdings durch die eine oder andere Art der Normierung in eine bestimmte Richtung gelenkt werden. Einige, wenn auch allgemeine, Richtlinien kann der Fachbereich seinen Lektoren schon mit auf den Weg geben. Alles andere wäre ein Ausdruck von geringem Respekt gegenüber Lektoren und Studenten. Wenn eindeutige Kriterien vorliegen (und eingehalten werden), wird es auch wesentlich einfacher, dem Studenten die gewünschte Stellungnahme zu seiner Leistung vorzulegen. Wenn die Studenten sehr unterschiedliche Rückmeldungen darüber erhalten, wie eine Aufgabe

zu beantworten ist, oder wo der Unterschied zwischen einer guten und weniger guten Antwort liegt, wird dadurch mehr Verunsicherung als Verdeutlichung bewirkt.

Ein weiterer wichtiger Punkt, auf den auch die beiden „Superprüfer“ eingingen, ist die Gestaltung der eigentlichen Prüfungsaufgaben. Einerseits geht es dabei um das sprachliche Leistungsniveau (z.B. Modifizierung des Verhaltens oder Verhaltensmodifizierung), andererseits aber auch um den Umfang der Arbeit, den die Studenten zu beantworten haben. Wie bereits erwähnt, stehen im Rahmen der Prüfung vier Stunden zur Verfügung. In einigen Fällen (PS 101 und PS 102) müssen die Studenten drei von vier Aufgaben in diesem Zeitraum lösen. Vor diesem Hintergrund ist es fast sinnlos, den Studenten z.B. folgende Aufgabe vorzulegen: „Über das Unbewusste“. Man kann auch nicht leicht nachzuvollziehen, wie ein Student im Verlauf dieser kurzen Zeit folgende Aufgabe lösen soll: „Lernpsychologische Grundsätze zur Änderung von Verhaltensweisen. Wodurch wird Verhalten geändert und beeinflusst?“ Vollständig unmöglich scheint folgende Aufgabe zu sein: „Kurze Abhandlung über die Entwicklung im Bereich der verhaltens- und gestaltpsychologischen Annäherung an die Psychologie. Gehen Sie auf die Gemeinsamkeiten und Unterschiede dieser beiden Perspektiven ein.“ Mir leuchtet es einfach nicht ein, wie manche Studenten (oder auch erfahrene Lektoren) eine solche Aufgabenstellung und noch zwei weitere im Verlauf dieser zur Verfügung stehenden Zeit mehr als oberflächlich lösen sollen. Wenn wir Studenten diese Art von Aufgaben in einer vierstündigen Prüfung stellen, wird sofort klar, dass die Prüfung kein Wissenstest ist, sondern eher andere Eigenschaften wie Stressbewältigung, Zeitplanung, Koordinationsleistung und Tempo testet. In den Prüfungsvorschriften für das Erstjahresstudium der Psychologie steht, dass die Prüfung keine Lernsituation ist. Daher drängt sich die Frage besonders auf, was die Prüfung darstellen soll, was also eventuell zu bewerten ist.

Obwohl unsere Ergebnisse keine Grundlage für eine Beurteilung dessen ist, inwiefern die Lektoren fachlich auf einem aktuellen Stand sind oder wie gut sie den Studienstoff beherrschen, liegen uns eindeutige Erkenntnisse vor, dass die Noten von der jeweiligen Kommission abhängen. Dies ist an und für sich keine Überraschung, weil Bewertungen immer unter einem subjektiven Einfluss stehen. Nicht zuletzt weisen auch die Anmerkungen der vier „Superprüfer“ darauf hin, dass solche subjektiven Beurteilungen eine Rolle spielen. Beunruhigend sind in diesem Zusammenhang jedoch die relativ großen Abweichungen in Einzelfällen und der Umstand, dass Hundertstel so entscheidend sind. Worauf die großen Unterschiede zurückzuführen sind, bleibt Spekulation. Dabei kann es sich selbstverständlich um eine Folge der Tatsache handeln, dass die Lektoren verschiedene Dinge für wichtig halten (dieser Eindruck entsteht, wenn man die Kommentare von „Superprüfer“ 2 liest). Es ist auch möglich, dass diese Unterschiede damit zusammenhängen, dass verschiedene Lektoren unterschiedlich viel über den Unterrichtsstoff wissen und den vorgenommenen Benotungen somit verschiedene Kriterien zugrunde liegen.

In diesem Zusammenhang muss darauf hingewiesen werden, dass die Lektoren nicht eine bestimmte Eigenschaft oder Leistung messen (als wäre dies überhaupt möglich), sondern vielmehr verschiedene Leistungen beurteilen sollen. Warum eine bestimmte Leistung mit 2,2 oder 2,4 be-

wertet wird, ist zunächst auch beliebig und daher unwesentlich. (Das trifft selbstverständlich nicht zu, wenn die Benotungen zwischen „Gut“ und „Durchgefallen“ schwanken.) Derartige Bewertungen leiden wie gesagt immer unter einer mehr oder weniger ausgeprägten Subjektivität. Aber erneut stehen wir beim Erstjahrestudium der Psychologie vor einer besonderen Situation, weil Zehntel (und Hundertstel) von so entscheidender Bedeutung sind. Wenn man eine Praxis beibehält, die auf derartig geringen Unterschieden basiert, darf man sich auch über entsprechenden die Signale an die Umwelt keine Illusionen machen, dass nämlich bestimmte Leistungen genau gemessen werden können und man Unterschiede zwischen den Noten 2,11 und 2,12 machen kann. Damit wird auf zahlreiche Arten auch zur Stärkung der verbreiteten Fehleinschätzung beigetragen, dass ein absoluter, korrekter und objektiver Ausdruck (Note) für eine bestimmte Antwort vorliegt. Dies ist keine Sondersituation für das Erstjahrestudium der Psychologie, sondern eher ein allgemeines Missverständnis. Wenn man diese falschen Voraussetzungen mit dem Gedanken verknüpft, dass ein bestimmtes Prüfungsergebnis (Note) eine absolute Aussage über das Potential und die zukünftige Entwicklung einer Person darstellt, wird es ganz verrückt, was auch Gardner (1999) zum Ausdruck bringt: „Wenn das Leben nur aus Unterricht bestünde, würden die meisten formellen Prüfungen ihren Zweck gut erfüllen, allerdings könnte man dann die Noten des Vorjahres genauso gut zum selben Zweck heranziehen. Unterricht ist jedoch vielmehr eine Vorbereitung auf das Leben, und es gibt wenig Grund zu der Annahme, dass formelle Prüfungen allein eine indifferente Aussage über den Erfolg ermöglichen, den ein Student nach der Schule zu erwarten hat“ (S. 97).

Gardner steht der traditionellen Prüfungspraxis grundsätzlich skeptisch gegenüber, weil diese so sehr Personen bevorzugen, die über eine hohe verbale und logisch-mathematische Intelligenz verfügen. Stattdessen schlägt er eine Praxis vor, bei der die Personen im Zuge des Lernprozesses bewertet werden, und zwar in natürlichen Situationen des Unterrichts („assessment-in-context“). In einer Erörterung der verschiedenen Alternativen sagt er: „Ich sehe als Haupthindernis für diese „kontextabhängige Bewertung“ nicht die Verfügbarkeit von Ressourcen, sondern den Mangel an gutem Willen. In diesem Land gibt es gegenwärtig den enormen Wunsch, Ausbildung einheitlich zu gestalten, alle Studenten gleich zu behandeln und dieselbe Art von eindimensionalen Messwerten für alle zu verwenden. Dieser Trend ist auf wissenschaftlicher Grundlage unangemessen und aus ethischen Gründen abzulehnen. Die gegenwärtige Geisteshaltung basiert teilweise auf einer verständlichen Abneigung gegenüber den Exzessen einiger früherer Bildungsexperimente, aber in einem ärgerlichen Ausmaß auch auf einer allgemeinen Feindseligkeit gegenüber Studenten, Lehrern und dem Lernprozess insgesamt“ (Gardner 1999, S. 115).

In einer Erörterung der Praxis im britischen Ausbildungssystem betont auch Kelly (1999) die Bedeutung der Unterscheidung zwischen Messung und Beurteilung: „... diese Praxis hat die Rückkehr zur überholten Betrachtungsweise auf ihre Fahnen geschrieben, bei der die Beurteilung als Form einer Messung und nicht als wesentlicher Beurteilungsprozess, der sie de facto ist, betrachtet wird. Der Begriff „Messung“ wird mit Genauigkeit und Präzision verbunden, aber bei einer näheren Betrachtung erkennt man so-

fort, dass bei den meisten Bildungsbeurteilungen wenige Genauigkeit vorzufinden ist. Und das Ausmaß der Genauigkeit und Präzision schwankt umgekehrt wiederum in Relation zur Komplexität und Perfektion des Beurteilungsobjekts. Genauigkeit ist einfacher zu erreichen, wenn wir relativ simple Formen von Errungenschaften beurteilen, aber auch dann ist sie wahrscheinlich nicht fehlerfrei“ (S. 128-129). Wenn wie im Fall des Erstjahrestudiums der Psychologie in Bergen keine Kriterien vorliegen, die man bei der Beurteilung der Antworten der Studenten heranziehen kann, ist die Arbeit des Prüfers natürlich sehr schwierig. Aber vielleicht noch wichtiger ist dabei der Umstand, dass der Fachbereich irgendwie die Verantwortung für die vorgenommenen Beurteilungen von sich weist und der einzelne Student dem Lektor überlassen wird. Somit wird die Frage, auf welchem aktuellen Stand der Lehre im Fachbereich der Lektor sich befindet (Lehrpläne, Stoffumfang und Vorlesungen), in der Realität ausgesprochen wichtig. Da auf Seiten des Fachbereichs die eindeutige Maßgabe besteht, dass keine Prüferanweisung erstellt wird, könnte man zu der Auffassung gelangen, dass gleichzeitig der Qualitätssicherung bei den Lektoren ein hoher Stellenwert beigemessen wird. Allerdings deutet wenig darauf hin, dass dies beim Erstjahrestudium der Psychologie an der Universität von Bergen der Fall ist. Die hauptsächliche Absicht bei dieser Arbeit war es nicht, die eigentliche Prüfform im Erstjahrestudium der Psychologie zu erörtern und alternative Beurteilungsformen vorzustellen. Jedoch spricht vieles dafür, dass die Frage der Interrater-Reliabilität in Verbindung mit der Prüf-/Bewertungsform zu betrachten ist. Zur Verteidigung einer Ordnung, bei der unzählige Studenten Jahre ihres Lebens mit der mehr oder weniger verzweifelten Jagd nach der „richtigen“ Note verbringen, wobei Hundertstel über „Erwählte“ und „Aussortierte“ entscheiden, lassen sich nur wenige Argumente anführen.

Literaturverzeichnis

- Engvik, H./Kvale, S./Havik, O. E. (1970): Rater reliability in evaluation of essay and oral examinations. *Scandinavian Journal of Educational Research*, 14, pp. 195-220.
- Fleming, N. D. (1999): Biases in marking students' written work: Quality? In: Brown, S./Glasner, A. (Eds.): *Assessment matters in higher education. Choosing and using diverse approaches*. Buckingham, pp. 83-92.
- Gabb, R. (1980): How consistent is examination marking? *Educational Services Unit. Newsletter no. 2*. New Zealand.
- Gardner, H. (1999): *Assessment in context*. In: Murphy, P. (Ed.): *Learners, learning & assessment*. London.
- Havik, O. E./Teigen, K. H./Kvale, S. (1969): *Endel analyse av karaktergivning ved psykologisk institutt*. Stensil. Psykologisk institutt, Universitetet i Oslo.
- Kelly, A. V. (1999): *The curriculum. Theory and practice*. London.
- Lundal, E./Manger, T./Mæland, K. (1991): *Rapport fra forsøk med flervalgsprøve. Psykologi grunnfag våren 1991*. Bergen.
- McDowell, L./Sambell, K. (1999): The experience of innovative assessment: student perspectives. In: Brown, S./Glasner, A. (Eds.): *Assessment matters in higher education. Choosing and using diverse approaches*. Buckingham, pp. 71-82.
- Raaheim, A. (1998): *Grunnfag psykologi ved Universitetet i Bergen. Noen tanker om kontroll, kvalitet, læring og eksamen*. (PLF-rapport nr. 1). Bergen.
- Raaheim, A. (1999): *Learning, knowledge, and assessment at university. Using assignment essays as an alternative to traditional exams*. København: Nordic Council of Ministers (in press).
- Rowntree, D. (1987): *Assessing students: How shall we know them?* London.

■ Dr. Arild Raaheim, Professor für Psychologie, Studienprogrammleiter, Studienprogramm für Universitätspädagogik, Universität Bergen/Norwegen, E-Mail: Arild.Raaheim@iuh.uib.no

Oskar Frischenschlager, Lukas Mitterauer, Irina Scharinger & Gerald Haidinger

Einfluss von Lernverhalten und Lernstrategien auf die paradoxen Geschlechtsunterschiede in den Erfolgsraten im Wiener Medizinstudium



Oskar Frischenschlager



Lukas Mitterauer



Irina Scharinger



Gerald Haidinger

Mit dem Studienjahr 2002/03 wurde an der Medizinischen Universität Wien ein reformiertes Curriculum Humanmedizin eingeführt, bei dem eine selektive Prüfung am Ende des ersten Studienjahres bestanden werden muss, um in das Zweite aufgenommen zu werden. Unsere bisherigen Forschungen haben an mehreren Jahrgängen übereinstimmend vier Erfolgsprädiktoren erbracht, darunter das männliche Geschlecht. In der vorliegenden prospektiven Studie des Beginner-Jahrgangs 2004/05 untersuchen wir den Einfluss von Lernstrategien und Lernverhalten auf den Studienerfolg in Abhängigkeit vom Geschlecht. Wir zeigen, dass der Prüfungserfolg im Bereich des Lernverhaltens gut präzidiert werden kann, jedoch weniger durch differenzierte Erfassung der Lernstrategien. Die unterschiedlichen Erfolgsraten nach dem Geschlecht können weder durch Lernverhalten noch durch Lernstrategien erklärt werden.

1. Einleitung

Das mit dem Studienjahr 2002/03 an der Medizinischen Universität Wien eingeführte reformierte Curriculum brachte neben vielen anderen Veränderungen auch eine deutliche Verringerung der Prüfungsereignisse mit sich. Am Ende jedes Studienjahres ist eine summative integrative Prüfung vorgesehen (SIP). Das Bestehen der SIP-1 (am Ende des ersten Studienjahres) ist Voraussetzung für die Zulassung in das zweite Studienjahr. Aufgrund der Gesetzeslage in Österreich, die einen weitgehend offenen Zugang zu den meisten Hochschulstudien vorsieht, stellte die SIP-1 das einzige Instrument der Medizinischen Universität dar, den Zugang zum weiteren Studium entsprechend der vorhandenen Kapazität zu regulieren. Von den mehr als 1.500 Studienanfängern wurden nur 660 in das zweite Studienjahr aufgenommen; Anlass genug, zu untersuchen, welche Faktoren für ein erfolgreiches Bestehen der SIP-1 prädiktiv sind.

In the study year 2002/03 a new medical curriculum was implemented at the Vienna Medical University. At the end of the first study year an exam must be taken deciding whether the student is allowed to enter the second study year or not. The previous studies yielded four predictors of success, one of them being male sex. In the present prospective study of the cohort of 2004/05 the authors investigated the influence of learning strategies and of learning behaviour on study success with respect to sex. The results of these studies are summarized by Oskar Frischenschlager, Lukas Mitterauer, Irina Scharinger and Gerald Haidinger in the report **The influence of learning strategies and learning behaviour on the unexpected sex-effect in study success at the Medical University of Vienna**. Their results render study behaviour being more predictive than highly differentiated aspects of learning strategies. The differences in study success by sex can neither be explained by learning behaviour nor by learning strategies.

In prospektiven Studien an den Jahrgängen 2002/03 und 2003/04 fanden wir vier Erfolgsprädiktoren: gute Schulleistungen, deutsche Muttersprache, hohe Lernkapazität und männliches Geschlecht (Frischenschlager et al 2005, Haidinger et al 2005). Auch konnte die Reliabilität der Befunde bestätigt werden (Haidinger et al 2006). Die deutlich geringere Bestehensrate weiblicher Studierender überrascht insofern, als

- a) deren Schulleistungen in der Regel besser sind, was auch in unseren Stichproben zutrif und dass
- b) die Schulleistungen generell den Studienerfolg präzisieren, was auch in unseren bisherigen Studien im Hinblick auf die SIP bestätigt wurde.

Dass der von uns gefundene Geschlechtseffekt eine Benachteiligung weiblicher Studierender darstellt, ergibt sich auch aus dem Befund, wonach ein Großteil jener weiblichen Studierenden, die die SIP-1 nicht rechtzeitig bestanden, mit einem Jahr Verlust ihr Studium fortsetzten (Mitterauer et al 2007). Wir folgern daraus die Annahme, dass die Erklärung dieses Geschlechtseffekts eher nicht im Bereich stabiler Variablen (trait) zu suchen ist, sondern eher bei veränderbaren Faktor (state), die durch Anpassungsleistungen verändert werden können.

Das Ziel der vorliegenden Studie ist, zu untersuchen, ob es Unterschiede im Lernverhalten bzw. in Lernstrategien gibt, die den geschlechtsspezifischen Erfolg bei der SIP-1 erklären können.

2. Methoden

Sechs bis drei Wochen vor dem ersten Termin der SIP-1 des Jahrgangs 2004/05 (Ende Juni 2005) wurden die Studierenden in Lehrveranstaltungen mit Anwesenheitspflicht gebeten, einen Fragebogen auszufüllen. Weiters wurde ihr schriftliches Einverständnis erbeten, die Fragebogendaten mit dem Prüfungsergebnis (Juni oder September), das im Internet veröffentlicht wird, in Beziehung zu setzen.

Neben den relevanten soziodemografischen Daten wurden die Schulleistungen, das aktuelle Lernverhalten (Zeitpunkt des geplanten Antretens, Einstellungen zum Prüfungssystem etc., 13 Items) erhoben. Lernstrategien wurden mittels einer Adaptierung des in zahlreichen Studien bewährten „Inventar zur Erfassung von Lernstrategien im Studium“ (LIST) von Wild et al (1992) bzw. Wild/Schiefele (1994) erhoben. Dieses erfasst nach einer Differenzierung von Kirby (1988) auf einer mittleren Generalisierungsebene zwischen Lernstilen und Lerntaktiken:

- a) kognitive Aspekte der Lernstrategien: Organisation (5 Items), Elaboration (9 Items), Wiederholen (4 Items),
- b) metakognitive Lernstrategien: Planung (7 Items), Überwachung (6 Items), Regulation (4 Items) und
- c) ressourcenbezogene Lernstrategien: Anstrengung (7 Items), Aufmerksamkeit (3 Items), Zeitmanagement (3 Items), Lernen mit Studienkollegen (4 Items), Gestaltung der Lernumgebung (2 Items, Literatur (2 Items).

Weiters wurde die Skala „Prüfungsangst“ (8 Items) aus dem Lernstrategieinventar „Wie lerne ich?“ (WLI) von Metzger et al 1994, die Skala „Subjektive Lernkompetenz“ (6 Items) von Cress (1999) sowie Cress und Friedrich (2001), die Skala „Erfolgservartung“ (7 Items) von Cress (1999) sowie Cress und Friedrich (2001) sowie selbst konstruierte Items (14 Items), die bereits in früheren Studien unserer Gruppe zur Erfassung der des SIP-spezifischen Lernverhaltens eingesetzt worden waren, verwendet. Insgesamt umfasste der Fragebogen 104 Items. Bei den Fragen zu den Lernstrategien, zur Prüfungsangst, Lernkompetenz, Erfolgservartung waren die Antworten auf einer Skala von 1 bis 6 (1: trifft gar nicht zu, 6: trifft völlig zu) zu geben.

Die Fragebögen wurden elektronisch erfasst und die Daten nach Veröffentlichung der Prüfungsergebnisse mit diesen verknüpft und mittels SPSS (Version 14) ausgewertet. Für dichotome Items wurden Chi-Quadrat-Tests, für intervall-

skalierte wurden mehrfaktorielle Varianzanalysen (Studien-erfolg und Geschlecht als abhängige Variablen) gerechnet. Als Signifikanzniveau wurde $P < 0,05$ angesetzt, für mehrfaches Testen wurde die Adjustierung nach Finner (1993) angewendet.

3. Ergebnisse

500 Studierende füllten den Fragebogen aus, davon war einer doppelt ausgefüllt, 11 waren schon einmal zur SIP-1 angetreten und mussten daher ausgeschlossen werden. Von den verbleibenden 488 traten 66 weder zum ersten noch zum zweiten Prüfungstermin an, sodass letztlich 422 Fragebögen zur Auswertung kamen.

Für die Repräsentativitätsprüfung der Stichprobe wurden die Geschlechtsverteilung und die Bestehensrate (Note bei SIP-1) der Stichprobe mit der Grundgesamtheit verglichen ($P = 0,158$). Die Stichprobe kann daher als repräsentativ angesehen werden.

In einem ersten Schritt wurde geprüft, ob sich die ursprüngliche Faktorenstruktur der Testinstrumente nach der Auswahl der am höchsten ladenden Items und der teilweisen situationsspezifischen Umformulierung von Items (Schiefele 2005) bestätigen lässt. „Insgesamt ließen sich die Originalfaktoren (Skalen) des LIST in dieser Stichprobe wiederfinden. Die Skalen „Kritisch Prüfen“ und „Zusammenhänge“ wurden jedoch aufgrund ihrer nur bedingt differenzierten Erfassung zur Skala „Elaboration“ zusammengefasst. Die Skala „Lernen mit Studienkollegen“ wurde trotz einer eher geringen Faktorenzugehörigkeit der Items in ihrer Ausgangsform beibehalten“ (Scharinger 2007, S. 92).

Nach Klärung dieser Frage wurde mittels mehrfaktorieller univariater Varianzanalysen untersucht, ob sich hinsichtlich der Faktoren/Skalen geschlechts- und erfolgsspezifische Unterschiede finden lassen und ob zwischen diesen Variablen Wechselwirkungen bestehen (Tabelle 1). Bei keiner der durchgeführten Varianzanalysen konnten signifikante Wechselwirkungen zwischen Geschlecht und Prüfungserfolg festgestellt werden.

Schulnoten: Die Summe der Schulnoten (Physik, Mathematik, Chemie, Englisch) hat einen signifikanten Einfluss auf den Prüfungserfolg ($P < 0,001$), je besser die Noten, desto häufiger wird die SIP bestanden. Männer weisen signifikant schlechtere Schulnoten als Frauen auf ($P = 0,002$). Der Durchschnitt der Schulnoten (Skala 1-5, 1=sehr gut, 5= nicht genügend) beträgt bei den Erfolgreichen 1,4, bei den nicht Erfolgreichen 2,1. weibliche Studierende weisen einen Notenschnitt von 1,8 auf, männliche 2,1.

Lernaufwand: Je größer der angegebene Lernaufwand, desto häufiger wird die SIP bestanden ($P = 0,006$). Zwischen Männern und Frauen findet sich kein signifikanter Unterschied. Im Gesamtdurchschnitt werden 22 Wochenstunden Lernaufwand neben den Lehrveranstaltungen angegeben, wobei Erfolgreiche angeben 26 Stunden zu lernen, nicht Erfolgreiche 21 Stunden.

Lernbeginn: Ein früherer Lernbeginn führt häufiger zum Bestehen der SIP ($P < 0,001$), ein Geschlechtsunterschied besteht nicht. Im Gesamtschnitt beginnen die Studierenden bereits im Jänner für die SIP zu lernen, 56% der Erfolgreichen, aber nur 29% der Erfolgrlosen lernen von Beginn an kontinuierlich mit.

Tabelle 1

	Erfolgreich		Erfolglos		Gesamt	
	N	MW	N	MW	MW	p
Noten gesamt	89	1,44	286	2,09	1,93	0,000
Lernaufwand	94	25,99	287	21,36	22,50	0,006
Lernbeginn	98	2,48	319	3,55	3,30	0,000
Punkte FIP	95	73,02	265	52,94	58,24	0,000
Organisation	98	3,38	307	3,46	3,44	0,853
Elaboration	93	4,01	294	3,89	3,92	0,470
Wiederholen	98	3,76	313	3,67	3,69	0,499
Planung	96	4,06	305	4,03	4,04	0,858
Überwachung	98	4,31	306	4,01	4,08	0,014
Regulation	97	3,98	308	4,08	4,06	0,560
Anstrengung	96	4,59	303	4,03	4,16	0,000
Aufmerksamkeit	96	3,97	309	3,75	3,80	0,272
Zeitmanagement	96	3,58	310	3,51	3,52	0,770
Lernen mit KollegInnen	98	3,79	313	3,52	3,58	0,065
Gestaltung						
Studienumgebung	98	5,05	312	4,78	4,85	0,056
Literatur	98	4,69	315	4,56	4,59	0,391
Prüfungsangst	97	3,03	297	3,37	3,29	0,065
Subjektive Lernkompetenz	97	4,14	299	3,66	3,77	0,000
Erfolgserwartung	96	3,76	308	3,37	3,46	0,003
	männlich		weiblich		Gesamt	
	N	MW	N	MW	MW	p
Noten gesamt	163	2,07	212	1,82	1,93	0,002
Lernaufwand	164	22,66	217	22,37	22,50	0,984
Lernbeginn	184	3,28	233	3,31	3,30	0,858
Punkte FIP	164	60,98	196	55,94	58,24	0,065
Organisation	176	3,27	229	3,57	3,44	0,084
Elaboration	168	4,03	219	3,83	3,92	0,190
Wiederholen	179	3,65	232	3,72	3,69	0,470
Planung	174	4,06	227	4,02	4,04	0,966
Überwachung	173	4,07	231	4,09	4,08	0,853
Regulation	175	4,04	230	4,07	4,06	0,897
Anstrengung	173	4,14	226	4,18	4,16	0,967
Aufmerksamkeit	178	3,88	227	3,74	3,80	0,694
Zeitmanagement	176	3,56	230	3,50	3,52	0,903
Lernen mit KollegInnen	179	3,50	232	3,65	3,58	0,384
Gestaltung						
Studienumgebung	179	4,70	231	4,97	4,85	0,030
Literatur	180	4,48	233	4,68	4,59	0,594
Prüfungsangst	175	2,95	219	3,56	3,29	0,000
Subjektive Lernkompetenz	172	3,92	224	3,66	3,77	0,065
Erfolgserwartung	175	3,73	229	3,25	3,46	0,000
	Erfolgreich		Erfolglos		Gesamt	
Risiko einen Block nicht zu bestehen	N	%	N	%	N	p
männlich	53	4	131	28	184	0,001
weiblich	47	15	190	31	237	0,065

Punkte bei der „FIP“ (formative integrative Prüfung am Ende des ersten Semesters, bei der Studierende antreten müssen, das Ergebnis dient ausschließlich der Selbstkontrolle): Je höher die erreichte Punktezahl desto eher wird die SIP bestanden (Erfolgreiche: 73 Punkte, Erfolgreiche 53 Punkte, $P < 0,001$), Geschlechtsunterschiede bestehen nicht.

LIST-Faktoren: kognitive Lernstrategien (Organisation, Elaboration, Wiederholen) zeigten hinsichtlich des Prüfungserfolgs keine Unterschiede; auch zwischen den Geschlechtern traten keine Unterschiede auf.

Bei den metakognitiven Lernstrategien (Planung, Überwachung, Regulation) zeigte nur die Überwachung einen positiven Zusammenhang mit dem Prüfungserfolg (Erfolgreiche 4,3, Erfolgreiche 4,0, $P = 0,014$), Männer und Frauen zeigen keine Unterschiede. Bei den ressourcenbezogenen Lernstrategien (Anstrengung, Aufmerksamkeit, Zeitmanagement, Lernen mit Studienkollegen, Gestaltung der Lernumgebung, Literatur) gibt es einen signifikanten Unterschied beim Prüfungserfolg: je höher die angegebene Anstrengung, desto eher erfolgt ein Bestehen der SIP (Erfolgreiche: 4,6, Erfolgreiche: 4,0; $P < 0,001$), in der Gestaltung der Studienumgebung gaben Frauen an, darauf mehr Bedacht zu nehmen (weiblich: 5,0, männlich: 4,7; $P = 0,030$).

Prüfungsangst: Die Prüfungsangst hat keinen Einfluss auf den Studienerfolg, Frauen weisen jedoch signifikant höhere Angstwerte auf als Männer (weiblich: 3,6, männlich: 3,0; $P < 0,001$).

Subjektive Lernkompetenz: Je höher die diesbezügliche Selbsteinschätzung, desto eher wird die SIP bestanden (Erfolgreiche: 4,1, Erfolgreiche: 3,7; $P < 0,001$), die Geschlechter unterschieden sich auch hier nicht.

Bewusst eingegangenes Risiko, einen Block nicht zu bestehen: (wer nur in einem der 6 Themenblöcke das Punktelimit nicht erreicht, muss nicht die gesamte SIP, sondern nur die Prüfung über den entsprechenden Block wiederholen) Männer, die dies riskieren, bestehen die SIP eher nicht (männlich-erfolgreich: 4%, männlich-nicht erfolgreich: 28%; $P = 0,001$), bei Frauen findet sich kein signifikanter Unterschied.

Erfolgserwartung: Bei der SIP erfolgreiche Studierende weisen eine höhere subjektive Erfolgserwartung auf (Erfolgreiche: 3,7, nicht Erfolgreiche: 3,3; $P = 0,003$), männliche Studierende weisen signifikant höhere Werte auf (männlich: 3,7, weiblich: 3,3; $P < 0,001$).

4. Diskussion

Die vorliegende Studie wurde unternommen, um die Paradoxie, wonach weibliche Studierende generell (sowie auch in unseren bisherigen Untersuchungen) bessere Schulleistungen, jedoch durchwegs deutlich schlechtere Durchkommensraten bei der SIP aufweisen, aufzuklären. Unsere bisherigen Arbeiten dazu führten zu der Vermutung, dass sich einerseits Lernstrategien auf das Bestehen der Prüfung auswirken und dass darüber hinaus ein Unterschied in den Lernstrategien zwischen den Geschlechtern besteht. Insbesondere weil im Rahmen der SIP sehr große Stoffmengen (Lernstoff eines ganzen Studienjahres) bewältigt werden müssen, vermuteten wir, dass den Lernstrategien eine besondere Relevanz zukommt. In der Planung unserer Studie haben wir daher versucht, eine breite Palette von Lernstrategien zu erfassen.

Unsere Ergebnisse bestätigen diese Annahme, wobei nicht so sehr die Art, an einen Lernstoff heran zu gehen, relevant scheint (wie dies der LIST erfasst), sondern vielmehr die

Lernintensität, die – wissend um die zu bewältigenden Stoffmengen – gekennzeichnet ist durch früheren Lernbeginn und eine größere Zahl an Wochenstunden, die für das Lernen aufgewandt werden. In diesen Befund fügt es sich sinnvoll ein, dass die Punktezahl bei der FIP (nach dem ersten Semester) ein weiterer Erfolgsprädiktor ist, indiziert dies doch ein frühes Ernstnehmen des Lernaufwandes. Auch die Ergebnisse bei den Schulleistungen, der subjektiven Erfolgserwartung und der subjektiven Lernkompetenz zeigen konsequenterweise in dieselbe Richtung, sind sie doch Ausdruck eines bisher erfolgreich eingesetzten Lernverhaltens, welches im Falle des Misserfolges (Nicht-Bestehen der SIP) geändert werden muss. Es scheint daher, dass das vom LIST erfasste und unter dem Konzept „Lernstrategien“ firmierende Faktorenbündel nicht wirklich jene Herangehensweisen erfasst, die für das Bewältigen sehr großer Stoffmengen erforderlich sind, sondern eher geeignet ist, den Lernvorgang bzw. die Lernaktivität „mikroskopisch“, d.h. im Detail aufzuschlüsseln. Die Tatsache, dass ein bestimmtes Lernverhalten (gute bisherige Lernkarriere, früher Beginn, laufend Mitlernen, Fleiß) günstig für das Bestehen der SIP ist, ist also eher ein „makroskopischer“ Befund. Er sagt nichts darüber aus, wie die Studierenden im Detail an die etwa 8.000 Seiten Stoff herangehen, um sich auf die Prüfung vorzubereiten. Vielleicht ist es also auch kein Zufall, dass der Geschlechtseffekt, der ebenfalls unabhängig besteht (und nun im vierten Studienjahr repliziert wurde) nicht mit den erhobenen Lernstrategien verknüpft ist. Mit anderen Worten: die erfolgsrelevanten Faktoren des Lernverhaltens zeigten keine geschlechtsspezifischen Unterschiede, andere Faktoren hingegen (Gestaltung der Studiumgebung, Prüfungsangst) zeigten deutliche Geschlechtsunterschiede, erwiesen sich aber nicht als erfolgsrelevant. Hinsichtlich der Schulleistungen konnte der Effekt repliziert werden, wonach Mädchen signifikant bessere Schulnoten aufweisen, die für sich ein signifikanter Erfolgsprädiktor sind, jedoch die SIP-1 signifikant weniger häufig bestehen. Die Erklärung für diesen paradoxen Geschlechtseffekt ist also vermutlich in einem anderen Themenbereich als dem der „mikroskopischen“ Lernstrategien oder des „makroskopischen“ Lernverhaltens zu suchen, soweit diese in unserer Studie erfasst wurden.

Möglicherweise könnten Erfahrungswerte mit Frauen in technischen Studien einen Hinweis liefern: Studentinnen beginnen eine technische Ausbildung mit anderen Voraussetzungen als ihre Kollegen. Ein Bund-Länder-Modellversuch „Frauen im Ingenieurstudium an Fachhochschulen, 1994-1997 an der Fachhochschule Bielefeld“ ergab, dass erworbene Unterschiede zwischen den Geschlechtern in Selbstbewusstsein, Durchsetzungsvermögen und Vertrauen in die eigene Leistungsstärke in Technik und Naturwissenschaft im Rahmen der Fortbildung von Dozentinnen und Dozenten sowie der Laboringenieurinnen und -ingenieure thematisiert werden sollten (Schwarze 1998). Dieser Appell richtet sich also an die Lehrenden bzw. die Curriculausplanung, sofern ein Geschlechtergleichgewicht angestrebt wird.

Interessant erscheinen darüber hinaus die Befunde zur Prüfungsangst, wonach Studentinnen deutlich mehr Angst angeben als Studenten; weiter gehende Schlüsse auf den Geschlechtseffekt verbieten sich allerdings, da der negative

Zusammenhang zwischen Prüfungsangst und Erfolg die Signifikanz knapp verfehlte.

5. Schlussfolgerungen

Was ist die Lehre aus diesem Versuch, den paradoxen Geschlechtseffekt in Lernstrategien bzw. im Lernverhalten aufzuspüren? Zum einen konnte der Effekt, wonach männliche Studierende gegen die Erwartung häufiger die SIP bestehen, erneut bestätigt werden. Ein Hinweis mehr also, dass es sich sicher nicht um einen Zufallsbefund handeln kann. Weiters eine Reihe von in sich schlüssigen Bestätigungen für die Erfahrung, dass der Stoff eines Studienjahres kontinuierliche Anstrengung und laufendes Mitlernen erfordert. Dass die ins Detail gehende, „mikroskopische“ Analyse des Lernverhaltens den Geschlechtseffekt nicht zu erklären vermag, überrascht aufs Erste, lässt aber die Interpretation zu, dass die grundsätzliche Herangehensweise an einen Lernstoff in diesem Fall nicht prüfungserfolgsrelevant ist.

In einer früheren Arbeit haben wir bereits die Vermutung geäußert, dass es eher der Stoffumfang als die Art des Stoffes ist, der bestimmter Strategien bedarf, um erfolgreich bewältigt zu werden und dass das Lernverhalten von Mädchen, das ihnen in der Sekundarschule (bzw. in anderen Studiengängen) zu guten, ja besseren Leistungen verhalf, in diesem Studium nicht adaptiv ist, weil es auf Detailwissen ausgerichtet, auf sicheres Bestehen angelegt ist und mit entsprechend intensivem, auf Vollständigkeit des Wissens ausgerichtetem Lernen einhergeht. Es entsteht der Eindruck, dass bei dem Versuch, möglichst viel Grundlagen in den ersten Studienabschnitt hinein zu packen, nicht bedacht wurde, dass der auf diese Weise zustande gekommene Stoffumfang seriös und in die Tiefe gehend nicht mehr zu bewältigen ist und dass gerade dadurch strategische, rein prüfungsorientierte Lerner im Vorteil sind. Es wird in weiteren Studien daher genau diese Frage zu prüfen sein, inwieweit strategisches, auf die spezifische Situation ausgerichtetes Lernen von männlichen Studierenden in höherem Maß angewandt wird und darüber hinaus wird von der Curriculausplanung zu überlegen sein, wie diesen „unerwünschten Wirkungen“ begegnet werden kann.

Literaturverzeichnis

- Cress, U. (1999): Personale und situative Einflussfaktoren auf das selbstgesteuerte Lernen Erwachsener. Regensburg.
- Cress, U./Friedrich H. F. (2000): Selbst gesteuertes Lernen Erwachsener. Eine Lerntypologie auf der Basis von Lernstrategien, Lernmotivation und Selbstkonzept. Zeitschrift für Pädagogische Psychologie 14, S. 193-204.
- Finner, H. (1993): On a monotonicity problem in Step-Down Multiple Test Procedures. J Am Stat Assoc 88, pp. 920-923.
- Frischenschlager, O./Haidinger, G./Mitterauer, L. (2005): Factors associated with academic success at Vienna Medical School: Prospective Survey. Croat Med J, Vol. 46/No. 1, pp. 58-65, <http://www.cmj.hr/2005/46/1/15726677.pdf>.
- Haidinger, G./Frischenschlager, O./Mitterauer, L. (2005): Prediction of success in the first-year-exam in the study of medicine - a prospective survey. Wiener Klinische Wochenschrift, 117 (23-24), S. 827-832.
- Haidinger, G./Frischenschlager, O./Mitterauer, L. (2006): Reliability of predictors of study success in medicine. Wiener Medizinische Wochenschrift, 156 (13-14), S. 416-20.
- Kirby, J. R. (1988): Style, strategy and skill in reading. In: Schmeck, R.R. (Ed.), Learning strategies and learning styles. New York, pp. 230-274.

- Metzger, C./Weinstein, C. E./Palmer, D. R. (1994): WLI- Hochschule, Wie lerne ich? Lernstrategieinventar für Studentinnen und Studenten. Arau.
- Mitterauer, L./Haidinger/G./Frischenschlager, O. (2007): Sex differences in study progress at Medical University of Vienna. GMS Zeitschrift für Medizinische Ausbildung, 24 (2), Doc111 (20070523). www.egms.de/de/journals/zma/2007-24/zma000405.shtml.
- Scharinger, I. (2007): Selbstreguliertes Lernen im Medizinstudium. Empirische Erfassung selbstregulierten Lernens im Zuge der ersten Summativen Integrierten Prüfung (SIP 1) im Studium der Humanmedizin an der Medizinischen Universität Wien (MUW). Diplomarbeit, Universität Wien.
- Schiefele, U. (2005): Prüfungsnaher Erfassung von Lernstrategien und deren Vorhersagewert für nachfolgende Lernleistungen. In: Artelt, C./Moschner, B. (Hg.): Lernstrategien und Metakognition. Implikationen für Forschung und Praxis. Münster, S. 13-41.
- Schwarze, B. (1998): Frauen und Technik - zwei konträre Welten?, In: Handbuch Hochschullehre, S. 1-20.
- Wild, K.-P./Schiefele U./Winteler, A. (1992): LIST. Ein Verfahren zur Erfassung von Lernstrategien. Gelbe Reihe: Arbeiten zur Empirischen Pädagogik und Pädagogischen Psychologie, S. 20.
- Wild, K.-P./Schiefele, U. (1994): Lernstrategien im Studium. Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. Zeitschrift für Differentielle und Diagnostische Psychologie, 15, S. 185-200.

- **Dr. Oskar Frischenschlager**, ao. Univ.-Professor für Medizinische Psychologie, Institut für Medizinische Psychologie des Zentrums für Public Health, Medizinische Universität Wien, E-Mail: oskar.frischenschlager@meduniwien.ac.at
- **Mag. Lukas Mitterauer**, Psychologe, Evaluator, Einrichtung für Qualitätssicherung, Universität Wien, E-Mail: lukas.mitterauer@univie.ac.at
- **Irina Scharinger**, M.A., Psychologin, E-Mail: a8300925@unet.univie.ac.at
- **Dr. Gerald Haidinger**, ao. Univ.-Professor für Sozialmedizin, Abteilung für Epidemiologie des Zentrums für Public Health, Medizinische Universität Wien, E-Mail: gerald.haidinger@meduniwien.ac.at

Sehr geehrte Leserinnen und Leser,

viele Hochschulen bauen zur Zeit ihre Tutorenprogramme aus. Im Folgenden bietet der UniversitätsVerlagWebler einen Titel die ideale Ratgeberliteratur für Tutorien.

Fachbereiche geben dieses Buch mit Hilfe der Studienbeiträge jedem Tutor als persönliche Ausstattung an die Hand („Verbrauchsmaterial“).

Das ist eine einmalige Investition in die Qualität der Tutorien, die den Studierenden direkt zugute kommt, also sich lohnt!

Helen Knauf: Tutorenhandbuch Einführung in die Tutorienarbeit

Das Tutorenhandbuch bietet eine grundlegende Einführung in die Tutorienarbeit und kann als Ideenschatz für die Gestaltung von Tutorien und Workshops zur Tutorienqualifizierung dienen. Einzelne Veranstaltungskonzepte laden zur Nachahmung ein; Erfahrungsberichte aus der Tutorienarbeit zeigen, wie Tutorien an Hochschulen etabliert werden können.

Das Handbuch gibt in übersichtlicher Form Antworten u.a. zu den Fragen:

- Was ist Tutorienarbeit?
- Wie kann erfolgreiche Tutorienarbeit geleistet werden?
- Welche Methoden finden Anwendung?

Für jeden, der sich mit Tutorienarbeit beschäftigt, ist dieses Buch ein unentbehrliches Arbeitsmittel.

*ISBN 3-937026-34-7, Bielefeld 2007,
2. überarbeitete Auflage, 159 Seiten, 22.80 Euro*

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



Förderprogramm „Hochschulforschung als Beitrag zur Professionalisierung der Hochschullehre“ des Bundesministeriums für Bildung und Forschung

Richtlinien zur Förderung von Forschung über Hochschulen im Rahmen des BMBF-Förderprogramms „Empirische Bildungsforschung“

1. Zuwendungszweck, Rechtsgrundlage

1.1 Zuwendungszweck

Kaum ein anderer gesellschaftlicher Teilbereich erfährt im Zuge des Strukturwandels hin zur Wissensgesellschaft einen größeren Bedeutungszuwachs als das Bildungssystem. Mit diesem Bedeutungszuwachs steigen gleichermaßen die gesellschaftlichen Erwartungen an dessen Leistungsfähigkeit wie an entsprechende Leistungsnachweise. Weitestgehend unstrittig ist dabei, dass es - neben ausreichender finanzieller Ressourcen - weiterer erheblicher Reform- und Modernisierungsanstrengungen mit Blick auf die Rahmenbedingungen sowie die inhaltliche und organisatorische Ausgestaltung der jeweiligen Bildungsbereiche bedarf, damit sich diese höheren Erwartungen auch erfüllen können.

Die einzelnen Bildungsbereiche sind von diesen Veränderungen in unterschiedlicher Weise betroffen.

Mit Blick auf die Hochschulen ist es zum einen der wissensbasierten Gesellschaften eigene Druck zur Höherqualifizierung, der dazu führt, dass Hochschulen inzwischen in weitaus größerem Umfang neben ihren Aufgaben im Bereich der Heranbildung des Nachwuchses für Wissenschaft und Forschung selbst auch gesellschaftliche Pflichten in der Berufsausbildung übernehmen. Die über die letzten Jahre kontinuierlich zunehmende Zahl neuer und meist anwendungsnaher Studiengänge ist Ausdruck dieser Entwicklung. Mit wachsender Notwendigkeit des lebenslangen Lernens sowie der für erfolgreiche Innovationsgesellschaften essenziellen Optimierung des Forschungstransfers gilt es zum anderen, die von Hochschulen bisher eher randständig wahrgenommene wissenschaftliche Weiterbildung bzw. das Verhältnis von Erstausbildung und Weiterbildung grundsätzlich neu zu justieren.

In einem deutlichen Missverhältnis zu der Wissensgesellschaften charakterisierenden Verwissenschaftlichung aller Lebens- und Arbeitsbereiche steht die Wissenschaft selbst. In Forschung und Lehre ist erst seit relativ kurzer Zeit ein gestiegenes Interesse an einer systematischen Beschäftigung mit dem Prozess der Generierung und Vermittlung von wissenschaftlichem Wissen, seinen Erfolgsvoraussetzungen - dazu gehört auch die Gestaltung der Institutionen - und Wirkungszusammenhängen zu verzeichnen. Die interne Gestaltung von Lehre und Forschung wie die externe Governance geschehen bislang wenig theoriegeleitet, die enormen in den letzten Jahren angestoßenen Modernisie-

rungs- und Reformmaßnahmen finden oft ohne adäquate wissenschaftliche Begleitung statt.

Insgesamt kann vermutet werden, dass eine stärkere Verwissenschaftlichung des Wissenschafts- bzw. Hochschulsystems selbst zu verbesserten Verhältnissen und Ergebnissen in Lehre und Forschung führt. Insbesondere kann erwartet werden, dass das den individuellen und kollektiven Akteuren an den Hochschulen sowie der Hochschulpolitik zur Verfügung stehende wissenschaftlich abgesicherte Wissen über die (un-)erwünschten Wirkungen ihres Handelns, dadurch vergrößert wird.

Übergeordneter Anspruch des neuen BMBF-Förderschwerpunkts Hochschulforschung ist es, diesem Defizit entgegen zu wirken, diesbezüglich vorhandene FuE-Kapazitäten zu stärken sowie - insbesondere - zusätzliche FuE-Kapazitäten aufzubauen und damit die Grundlage für objektivierbare Entscheidungen der Hochschulakteure zu vergrößern. Als ein genuin interdisziplinäres Forschungsfeld, in dem substantielles Gestaltungswissen nur durch disziplinübergreifende Forschungsarbeiten generiert werden kann, betrifft diese Stärkung der FuE-Basis eine ganze Reihe wissenschaftlicher Einzeldisziplinen.

1.2 Rechtsgrundlage

Vorhaben können nach Maßgabe dieser Richtlinien, der BMBF-Standardrichtlinien für Zuwendungen auf Ausgaben- bzw. Kostenbasis und der Verwaltungsvorschriften zu §44 Bundeshaushaltsordnung (BHO) durch Zuwendungen gefördert werden. Ein Rechtsanspruch auf Gewährung einer Zuwendung besteht nicht. Der Zuwendungsgeber entscheidet auf Grund seines pflichtgemäßen Ermessens im Rahmen der verfügbaren Haushaltsmittel.

2. Gegenstand der Förderung

Das erste im Rahmen des BMBF-Förderprogramms „Empirische Bildungsforschung“ speziell zum Hochschulbereich entwickelte Förderangebot zielt auf den Auf- bzw. Ausbau wissenschaftlicher Grundlagen der Hochschullehre.

Gefördert werden FuE-Projekte, die

- die curriculare Gestaltung (einschl. eines Systems der Leistungskontrolle) behandeln und dabei die veränderten an Hochschulen gerichteten Anforderungen (gestiegene Bedeutung der Kreativität, höhere Kompetenzorientierung, Differenzierung der Hochschullehre hinsichtlich ihrer Berufsorientierung ...) sowie die gestiegene Bedeutung der Lernorte außerhalb der Hochschulen (Praxis, informelles Lernen) berücksichtigen,

- die formal-organisatorische Gestaltung (einschl. eines Systems der Effizienz- und Qualitätsmessung) behandeln und dabei insbesondere Ansätze zur Professionalisierung der Lehrenden sowie der Support-Funktionen (Studienberatung, Studienbegleitung, CareerServices, neue Berufsmuster im Bereich des Hochschulmanagements ...) sowie zur Gestaltung der Schnittstelle zu den vor- bzw. nachfolgenden Bereichen Schule und Wirtschaft berücksichtigen,
- die Lehr-Lern-Prozessgestaltung behandeln und dabei insbesondere Ansätze zu dessen Individualisierung bearbeiten - sowohl mit Blick auf eine Erhöhung der Selbststeuerung und der höheren Bedeutung außerinstitutioneller Lernorte als auch der individuellen Förderung unter den Bedingungen gesteigerter Studierendenzahlen.

Die Ansätze sollen sich vornehmlich, aber nicht ausschließlich auf die wissenschaftliche Erstausbildung beziehen und die neuen technischen Möglichkeiten zur Modernisierung der Lehre sowie den im Zuge politischer Reformmaßnahmen in den letzten Jahren deutlich erweiterten organisatorisch-rechtlichen Gestaltungsspielraum der Hochschulen (z. B. durch die Bologna-Reform) nutzen.

Nicht förderfähig sind

- Projekte mit inkrementellem hochschuldidaktischen Anspruch ohne Berücksichtigung der organisatorisch-strukturellen Rahmenbedingungen,
- Projekte, in deren Mittelpunkt die Entwicklung von Bildungssoftware steht,
- Forschungsvorhaben mit reiner oder überwiegender Reflexionsorientierung ohne Anwendungsbezug.

3. Zuwendungsempfänger

Antragsberechtigt sind in erster Linie Hochschulen sowie außeruniversitäre - auch als Wirtschaftsunternehmen geführte - Forschungseinrichtungen.

Forschungseinrichtungen, die gemeinsam von Bund und Ländern grundfinanziert werden, kann nur unter bestimmten Voraussetzungen eine Projektförderung für ihren zusätzlichen Aufwand bewilligt werden.

4. Zuwendungsvoraussetzungen

- Förderinteressierte, die Projektvorschläge entsprechend der unter 2. beschriebenen inhaltlichen Anforderungen sowie des unter 7. dargestellten Verfahrens einreichen, sollten auf dem Gebiet der Hochschulforschung ausgewiesen sein. Angesichts der vergleichsweise noch in geringem Umfang vorhandenen FuE-Kapazitäten im Bereich der Hochschulforschung in Deutschland wird zusätzlich die Möglichkeit des Aufbaus eines Projektteams eingeräumt (vgl. 5.)
- Mit Blick auf die übergeordnete Zielsetzung des Förderangebots (s. unter 1.1) sind vorrangig solche Vorhaben förderfähig, die als konkrete Interventions- bzw. Begleitforschungsprojekte angelegt sind bzw. die ihre Nachhaltigkeit über bereits im Projekt verankerte Transfermaßnahmen sichern - z. B. durch die Struktur des Projekts als Verbundprojekt mit Universitäten und/oder Fachhochschulen als Erprobungspartner.

- Die Projektvorschläge sollen möglichst mehr als einen der unter 2. genannten drei zentralen Gestaltungsaspekte thematisieren. Entsprechend dieser Anforderung sollen die zur Förderung vorgeschlagenen Vorhaben vorrangig durch interdisziplinär zusammengesetzte Forschungsverbände bearbeitet werden.

Antragsteller sollen sich - auch im eigenen Interesse - im Umfeld des national beabsichtigten Vorhabens mit dem EU-Forschungsrahmenprogramm vertraut machen. Sie sollen prüfen, ob das beabsichtigte Vorhaben spezifische europäische Komponenten aufweist und damit eine abschließliche EU-Förderung möglich ist. Weiterhin ist zu prüfen, inwieweit im Umfeld des national beabsichtigten Vorhabens ergänzend ein Förderantrag bei der EU gestellt werden kann. Das Ergebnis der Prüfungen soll im nationalen Förderantrag kurz dargestellt werden.

Die Partner eines „Verbundprojekts“ haben Ihre Zusammenarbeit in einer Kooperationsvereinbarung zu regeln. Vor der Förderentscheidung muss eine grundsätzliche Übereinkunft über bestimmte vom BMBF vorgegebene Kriterien nachgewiesen werden.

Einzelheiten können einem BMBF-Merkblatt - Vordruck O110 unter:

<http://www.kp.dlr.de/profi/easy/bmbf/pdf/O110.pdf> entnommen werden.

5. Art und Umfang, Höhe der Zuwendung

Die Zuwendungen können im Wege der Projektförderung als nicht rückzahlbare Zuschüsse gewährt werden.

Bemessungsgrundlage für Zuwendungen an Unternehmen der gewerblichen Wirtschaft sind die zuwendungsfähigen projektbezogenen Kosten, die in der Regel - je nach Anwendungsnähe des Vorhabens - bis zu 50% anteilfinanziert werden können. Nach BMBF-Grundsätzen wird eine angemessene Eigenbeteiligung - grundsätzlich mindestens 50% der entstehenden zuwendungsfähigen Kosten - vorausgesetzt. Bemessungsgrundlage für Hochschulen, Forschungs- und Wissenschaftseinrichtungen und vergleichbare Institutionen sind die zuwendungsfähigen projektbezogenen Ausgaben (bei Helmholtz-Zentren und der Fraunhofer-Gesellschaft - FhG - die zuwendungsfähigen projektbezogenen Kosten), die individuell bis zu 100% gefördert werden können. Die Bemessung der jeweiligen Förderquote muss den Gemeinschaftsrahmen der EU-Kommission für staatliche FuE-Beihilfen berücksichtigen. Dieser Gemeinschaftsrahmen lässt für Kleine und Mittlere Unternehmen (KMU) eine differenzierte Bonusregelung zu, die ggf. zu einer höheren Förderquote führen kann. Zusätzlich wird im Rahmen dieser Förderbekanntmachung die Möglichkeit des geförderten Aufbaus eines Projektteams („Werkstattphase“) eingeräumt. Dieses Angebot richtet sich sowohl an Wissenschaftler und Wissenschaftlerinnen aus der Hochschulforschung als auch an solche aus „benachbarten Gebieten“ mit Interesse an diesem Förderangebot, die die zur Durchführung ihrer Projektidee notwendigen personellen Ressourcen erst entwickeln müssen. Eine solche Werkstattphase sollte 2 Jahre nicht überschreiten. Das Fördervolumen ist auf max. 200.000 Euro p.a. begrenzt.

6. Sonstige Zuwendungsbestimmungen

Bestandteil eines Zuwendungsbescheides auf Kostenbasis werden grundsätzlich die Allgemeinen Nebenbestimmungen für Zuwendungen auf Kostenbasis des BMBF an Unternehmen der gewerblichen Wirtschaft für FuE-Vorhaben (NKBF98). Bestandteil eines Zuwendungsbescheides auf Ausgabenbasis werden die Allgemeinen Nebenbestimmungen für Zuwendungen zur Projektförderung (ANBest-P) und die Besonderen Nebenbestimmungen für Zuwendungen des BMBF zur Projektförderung auf Ausgabenbasis (BN-Best-BMBF98).

7. Verfahren

7.1 Einschaltung eines Projektträgers und Anforderung von Unterlagen

Mit der technisch-administrativen Abwicklung der Fördermaßnahme hat das BMBF beauftragt:

DLR-Projektträger,
Stichwort „Hochschulforschung“,
Heinrich-Konen-Straße 1,
53277 Bonn

Ansprechpartnerin für inhaltliche Fragen:

Fr. Diegelmann

Ansprechpartner beim Projektträger:

Herr Michael Kindt

Vordrucke für Förderanträge, Richtlinien, Merkblätter, Hinweise und Nebenbestimmungen können unter der Internetadresse:

<http://www.kp-dlr.de/profi/easy/bmbf/index.htm>
abgerufen oder unmittelbar beim Projektträger angefordert werden.

Zur Erstellung von Projektskizzen wird die Nutzung des für die Bekanntmachung eingerichteten elektronischen Skizzentools dringend empfohlen:

<http://www.zukunftswerkstatt-hochschullehre.de>
Förmliche Förderanträge sind mit Hilfe des elektronischen Antragssystems „easy“ zu erstellen:
<http://www.kp.dlr.de/profi/easy/index.html>

7.2 Zweistufiges Förderverfahren

Das Förderverfahren ist zweistufig angelegt.

7.2.1 Vorlage und Auswahl von Projektskizzen

In der ersten Verfahrensstufe sind dem o.g. DLR-Projektträger bis spätestens 12.02.2008 zunächst Projektskizzen - möglichst in elektronischer Form über das o.g. Skizzentool und auf dem Postweg vorzulegen. Bei Verbundprojekten sind die Projektskizzen in Abstimmung mit dem vorgesehenen Verbundkoordinator vorzulegen. Die Vorlagefrist gilt nicht als Ausschlussfrist. Verspätet eingehende Projektskizzen können aber möglicherweise nicht mehr berücksichtigt werden. Projektskizzen sind wie folgt zu gliedern:

1. Titel des Vorhabens
2. Anschrift (einschl. Tel., Fax, E-Mail) (je Projektpartner)
3. Zielsetzung des Vorhabens (max. 2 Seiten)
4. Kurzbeschreibung des Vorhabens (max. 5 Seiten)
5. Angaben zur Erfüllung der Zuwendungsvoraussetzungen
6. Finanzierungsplan
7. ggf. Kurzbeschreibung der Notwendigkeit, der Ausgestaltung und des Finanzbedarfs für eine „Werkstattphase“

Weitere Unterpunkte entnehmen Sie bitte dem elektronischen Skizzentool. Aus der Vorlage einer Projektskizze kann ein Rechtsanspruch nicht abgeleitet werden.

Die eingegangenen Projektskizzen werden unter Beteiligung externer Gutachter/innen nach folgenden Kriterien bewertet:

- Erfüllung der Zuwendungsvoraussetzungen,
- Innovationshöhe/Originalität des Projektvorschlags,
- Relevanz für Qualität und Effizienz der Hochschullehre insgesamt,
- Erfolgsaussichten,
- Schlüssigkeit des Nachhaltigkeitskonzepts,
- Eignung des Standorts.

Zusätzlich wird der - langfristig - zu erwartende (direkte oder indirekte) Beitrag des vorgeschlagenen Ansatzes zur Bewältigung aktueller hochschulpolitischer Herausforderungen (Senkung der Studienabbruchquote, Steigerung des Studierendenanteils an MINT-Fächern, Verringerung der sozialen Selektion ...) betrachtet.

Auf der Grundlage der Bewertung werden dann die für eine Förderung geeigneten Projektideen ausgewählt. Das Auswahlergebnis wird den Interessenten schriftlich mitgeteilt.

7.2.2 Vorlage förmlicher Förderanträge und Entscheidungsverfahren

In der zweiten Verfahrensstufe werden die Interessenten bei positiv bewerteten Projektskizzen aufgefordert einen förmlichen Förderantrag vorzulegen, über den nach abschließender Prüfung entschieden wird. Für die Bewilligung, Auszahlung und Abrechnung der Zuwendung sowie für den Nachweis und die Prüfung der Verwendung und die ggf. erforderliche Aufhebung des Zuwendungsbescheides und die Rückforderung der gewährten Zuwendung gelten die Verwaltungsvorschriften zu §44 BHO sowie §§48 bis 49a Verwaltungsverfahrensgesetz (VwVfG), soweit nicht in diesen Förderrichtlinien Abweichungen zugelassen sind.

8. Inkrafttreten

Diese Förderrichtlinien treten mit dem Tag der Veröffentlichung im Bundesanzeiger in Kraft.

Quelle: <http://www.bmbf.de/foerderungen/11878.php>,
10.12.2007

Boris Schmidt



Boris Schmidt

Angenehm, konstruktiv – und nicht allzu wirkungsvoll? Lehrveranstaltungsevaluation aus der Sicht von Studierenden, Lehrenden und Evaluationsanbietern

Die Hochschulen im deutschsprachigen Raum befinden sich in einer Phase tiefgreifenden Wandels, besonders, aber nicht ausschließlich, im Bereich der Hochschullehre: Die Evaluation und Akkreditierung von Studiengängen (Benz/Kohler/Landfried 2004), Ansätze des Qualitätsmanagements der Lehre (Schmidt/Loßnitzer 2007) und die Intensivierung der Lehrveranstaltungsevaluation (Rindermann, 1997) sind nur einige Beispiele, an denen dieser Wandel konkret sichtbar wird. Die Veränderungsprozesse sind eingebettet in weitreichende gesellschaftliche und ökonomische Entwicklungen, die zu einem veränderten Verständnis des Bildungsauftrags und der Arbeitsweise der Hochschulen geführt haben: Potenzielle, aktuelle und ehemalige Studierende, die Politik und die Öffentlichkeit, Arbeitgeber in der Wirtschaft, der Industrie und im Bildungssektor und nicht zuletzt: auch die Menschen an der Hochschule selbst hinterfragen, in welcher Weise und mit welcher Qualität Hochschullehre stattfinden soll und welche (Evaluations-)Instrumente und Verfahren geeignet sind, um diese Qualität einzuschätzen, zu reflektieren und weiterzuentwickeln. Dieser Diskussionsprozess ist von einem Aufeinanderprallen unterschiedlichster Sichtweisen und Überzeugungen gekennzeichnet: Die jeweiligen Interessen und die individuellen Perspektiven der „Stakeholder“ (Daxner 1999, S. 45) spiegeln sich in den vorgebrachten Einschätzungen wider. Die Wahrnehmungen liegen dabei nicht selten so weit auseinander und gehen dabei zugleich von so unterschiedlichen Grundannahmen aus, dass man sich fragen kann, ob diejenigen, die dort streiten, überhaupt über dieselbe Hochschule, über ein und dasselbe Instrument, ein und dasselbe Verfahren sprechen oder über gänzlich Unterschiedliches.

1. Lehrveranstaltungsevaluation in der Diskussion

Ein prägnantes Beispiel, in dem sich die Unterschiedlichkeit der Perspektiven zeigt, ist die Lehrveranstaltungsevaluation. Lehrveranstaltungsevaluation (nachfolgend mit „LVE“ abgekürzt) ist eine spezifische Form der Lehrkritik,

In his comparing article **Agreeable, constructive – and not to effective? Evaluation of study courses from the viewpoint of students, tutors and evaluators** Boris Schmidt is asking for a methodical conception as well as for the effects of course evaluation.

die durch den „Einsatz standardisierter fragebogengestützter Studierendenbefragungen zu einzelnen Lehrveranstaltungen“ (Schmidt/Loßnitzer 2007, S. 48) gekennzeichnet ist. LVE unterscheidet sich durch die systematische Durchführung von anderen Formen der Lehrkritik wie beispielsweise offenen Feedback-Runden, adhoc-Fragebögen oder Lehrhospitationen. Innerhalb der LVE gibt es zahlreiche Differenzierungen, die sich unter anderem in den eingesetzten Instrumenten (z.B. Rindermann 1997, S. 20) niederschlagen. Ein Kernbereich der erfragten Aspekte (z.B. Strukturierung, Verständlichkeit, didaktische Aufbereitung) findet sich dennoch in nahezu allen Instrumenten wieder. Demgegenüber unterscheiden sich die Gestaltungsmerkmale bei der LVE-Durchführung und die angestrebten Wirkungsdimensionen zum Teil erheblich (vgl. Hage 1996; Daxner 1999): Wem werden die Evaluationsergebnisse zur Kenntnis gebracht (z.B. allen oder niemandem außer der Lehrperson), welche Rolle spielen die einzelnen Beteiligten bei der Durchführung (z.B. Dialogpartner oder „Messobjekte“), welche Anschlussmaßnahmen nach der Rückspiegelung der LVE-Ergebnisse gibt es (z.B. keine oder individuelle hochschuldidaktische Beratung)?

Welche Ausgestaltung dieser Durchführungsmerkmale als angemessen empfunden wird, hängt auch von Überzeugungen über LVE ab. Aleamoni (1987) formuliert sieben „typische Bedenken“, die häufig von denjenigen Lehrenden geäußert werden, deren Veranstaltungen im Fokus der Studierendeneinschätzung stehen: (1) Studierende können keine konsistenten Einschätzungen abgeben, (2) nur herausragende Fachwissenschaftler/-innen können Lehrqualität angemessen einschätzen, (3) LVE-Fragebögen erheben vor allem die Popularität der Lehrenden, (4) die Qualität der Lehre zeigt sich erst Jahre nach Verlassen der Hoch-

schule, (5) erwartete oder tatsächlich erzielte Noten schlagen sich maßgeblich in den LVE-Ergebnissen nieder, (6) zahlreiche externe Einflussgrößen verzerren die Studierendeneinschätzung, (7) LVE-Fragebögen sind generell nicht reliabel und nicht valide. Basierend auf einer Vielzahl von Untersuchungen kommt Aleamoni (1987) jedoch zu dem Schluss, dass die ersten fünf dieser Bedenken einer empirischen Überprüfung überhaupt nicht standhalten können, die verbleibenden zwei nur bedingt: Einzelne Einflussfaktoren wie zum Beispiel der Besuchsgrund (verpflichtend vs. freiwillig) korrelieren in einigen Untersuchungen in der Tat signifikant mit Studierendeneinschätzungen (Aleamoni 1987, S. 28f.; vgl. Rindermann 1997), und die eingesetzten LVE-Instrumente unterscheiden sich auch hinsichtlich ihrer Reliabilität und Validität – es gibt bessere und schlechtere, aussagekräftigere und weniger aussagekräftige Fragebögen (vgl. Marsh 1984). Unter dem Strich erweist sich jedoch die wissenschaftliche Grundlage, Lehrveranstaltungsevaluation an sich für aussagegelos zu halten und aus diesem Grund an ihrer Wirksamkeit zu zweifeln, als dürftig.

Forscherinnen und Forscher, die selber an derartigen Instrumenten zur Lehrveranstaltungsevaluation arbeiten oder LVE innerhalb ihrer eigenen Institution als Dienstleistung anbieten, kommen daher zu Einschätzungen, die sich nicht mit den genannten kritischen Haltungen vieler Lehrender decken. So resümiert Marsh im Rückblick auf 25 Jahre Forschung über Lehrveranstaltungsevaluation (1984, S. 707), dass sich LVE entgegen aller Bedenken als relativ valides und für die wesentlichen Zielsetzungen adäquates Verfahren erwiesen habe. Metaanalysen (z.B. Cohen 1980) geben zudem eine Reihe von konkreten Hinweisen auf diejenigen Durchführungsmerkmale, welche wirkungsvolle von wirkungsloser LVE unterscheiden (vgl. Hage 1996; Rindermann 1997): Der auf Basis der LVE-Ergebnisse stattfindende Dialog, die Auseinandersetzung mit den Ergebnissen, die Existenz und Nutzung von Angeboten zur lehrbezogenen Kompetenzentwicklung sind einige der Aspekte, die maßgeblich die Wirkung von LVE beeinflussen.

Die Perspektive der Studierenden als zentraler Interessengruppe bei der LVE wird nur in vereinzelt wissenschaftlichen Beiträgen zum Gegenstand der Betrachtung. Schweer/Rosemann (1995) untersuchten beispielsweise die Erwartungen an Lehrveranstaltungen, die den studentischen Einschätzungen zugrunde liegen, und wiesen eine hohe Heterogenität sowie signifikante Zusammenhänge mit den abgegebenen LVE-Einschätzungen nach. Rindermann (1999) ermittelte eine Reihe von Unterschieden zwischen Studierenden und Lehrenden in den Beschreibungen der Prozessmerkmale „guter Lehre“, die für den Umgang mit LVE ebenfalls von Relevanz sind. Aus Sicht der Studierenden wird Lehr(veranstaltungs)evaluation zudem nicht nur als Beitrag zum Qualitätsmanagement, sondern auch und vor allem als Instrument der Beteiligung und Mitbestimmung verstanden, denn die Studierenden seien „eben keine Kunden der Hochschule, sondern ihre Mitglieder. Sie haben eine aktive, mitgestaltende Rolle“ (Schulze 2001, S. 17).

Entsprechend untersucht auch die ESIB, der Verband der europäischen Studierendenvertretungen, Lehrveranstaltungsevaluation als Teilaspekt der Studierendenbeteiligung (ESIB 2003, S. 32ff.). Angesichts dieser Vielfalt und Unterschiedlichkeit zwischen den Perspektiven setzt sich die Diskussion fort, ob und in-

wieweit Lehrveranstaltungsevaluation ein geeignetes Instrument zur Lehrentwicklung sei. Rindermann (2000) hat diese Perspektivenunterschiede unter dem Stichwort des „Selbstobjektivierungsproblems“ an Hochschulen gezielt untersucht. Anhand von N=149 einschlägigen Publikationen im deutschsprachigen Raum wurde der Zusammenhang zwischen der Perspektive des Autors bzw. der Autorin und den getroffenen Aussagen über LVE untersucht.

Die Mehrheit der von Professoren/innen stammenden Veröffentlichungen gelangt zu dem Schluss, dass studentische Lehrveranstaltungsevaluation nicht oder nur bedingt valide sei (41,8% bzw. 34,5% der Publikationen), während Angehörige des Mittelbaus, Studierende sowie Hochschulexterne in ihren Beiträgen mehrheitlich zum Ergebnis gelangen, dass LVE sehr wohl die notwendige Validität aufweise (72,9%, 83,3% bzw. 58,8% der Publikationen). Bei der Analyse der Durchführungsmerkmale zeigte sich eine ähnliche Diskrepanz: Nur 20,0% der Professoren/innen stellen in ihren Publikationen eine über bloßes Feedback hinausgehende LVE-Durchführung (z.B. persönliche Ergebnisvorstellung) als wirkungsvoll heraus. Bei den Studien aus den Reihen der Mittelbauangehörigen liegt dieser Anteil bei 32,3%, und Studierende sowie Externe gelangen mehrheitlich zur Erkenntnis, dass eine entsprechend intensivierte Durchführung der LVE wirkungsvoll sei (66,7% bzw. 64,7%).

Woher rühren diese unterschiedlichen Einschätzungen? Ein Grund könnte in der Tendenz zur „Symmetrie von objektiver und subjektiver Wirklichkeit“ (Rindermann 2000, S. 78) liegen.

Die unterschiedlichen Wahrnehmungen der „Stakeholder“ über die Eignung von LVE-Instrumenten, und ihre Annahmen über Zusammenhänge zwischen LVE-Durchführungsmerkmalen und deren Wirkungen werden demnach sowohl vom individuell objektiv wahrnehmbaren Ausschnitt der Wirklichkeit als auch von subjektiven Faktoren wie der ungleich verteilten Zuschreibung von Expertise und von bestehenden Machtverhältnissen beeinflusst. Zusammen ergibt sich hieraus eine bestimmte Wirklichkeitskonstruktion über LVE. Angesichts dieser Wechselwirkungen wäre es eher der unwahrscheinliche als der wahrscheinliche Fall, wenn sich die Wahrnehmungen über das Instrument „Lehrveranstaltungsevaluation“ zwischen den Perspektiven decken würden.

2. Ziele der Untersuchung

In der vorliegenden Untersuchung wird diese potenzielle Unterschiedlichkeit bei drei ausgewählten, für die Gestaltung der LVE wesentlichen Perspektiven näher beleuchtet. Im Fokus stehen (1) an LVE teilnehmende Studierende, (2) teilnehmende Lehrende, (3) Anbieter von LVE, d.h. Personen, die den LVE-Prozess organisieren und gestalten. Gegenstand der Analyse sind die Durchführungsmerkmale und ausgewählte Wirkungsdimensionen sowie der Zusammenhang zwischen den Durchführungsmerkmalen einerseits und den wahrgenommenen Wirkungen andererseits. Durch die Untersuchung dieses Zusammenhangs soll auch die subjektive Wirklichkeit der Beteiligten im Sinne der sich in ihren Einschätzungen niederschlagenden Erklärungsmuster in den Blick genommen werden. Drei Fragestellungen werden behandelt:

- (1) Wie werden (1a) Durchführungsmerkmale sowie (1b) Wirkungsdimensionen der LVE vor dem Hintergrund individueller Erfahrungen mit Lehrveranstaltungsevaluation eingeschätzt?
- (2) Inwieweit unterscheiden sich die Wahrnehmungen zur Durchführung sowie zur Wirkung von LVE zwischen Studierenden, Lehrenden und LVE-Anbietern?
- (3) Weisen Studierende, Lehrende und LVE-Anbieter unterschiedliche Erklärungsmuster über den Zusammenhang zwischen Durchführungsmerkmalen und Wirkungsdimensionen von LVE auf?

Dabei soll nicht ein spezifisches LVE-Instrument und auch nicht ein bestimmtes Modell der LVE-Durchführung untersucht werden. Vielmehr wird der Versuch unternommen, in einem breiten Überblick über zahlreiche unterschiedliche Varianten und Organisationsformen an einer Vielzahl von Hochschulen zu Aussagen über „die“ Lehrveranstaltungsevaluation als einem aktuellen Ansatz zur Qualitätsentwicklung der Lehre zu gelangen.

3. Methode

Zum Einsatz kam ein insgesamt sechs Themengebiete umfassender Online-Fragebogen, der auf den Forschungsseiten des Universitätsprojekts Lehrevaluation der Friedrich-Schiller-Universität Jena eingerichtet und zwischen Mai und Juli 2006 freigeschaltet war. Zur Erhebung der individuell beobachteten LVE-Durchführungsmerkmale wurden insgesamt neun typische Aspekte ausgewählt, hinsichtlich derer sich konkrete Ansätze der Lehrveranstaltungsevaluation auch unabhängig von den jeweils eingesetzten Fragebögen unterscheiden können (vgl. Tabelle 1, Aspekte 1 bis 9). Auf Seiten der wahrgenommenen Wirkungen wurden drei Wirkungsdimensionen herausgegriffen, die potenzielle Ergebnisse der LVE-Durchführung beschreiben, nämlich dass Studierende sowie Lehrende von der Durchführung profitieren und dass es durch LVE zu spürbaren Verbesserungen der Lehrveranstaltungen kommt (vgl. Tabelle 1, Aspekte 10 bis 12).

Aus der Gesamtliste der deutschen Hochschulen wurde eine Stichprobe von N=53 Universitäten und Fachhochschulen durch Zufallsauswahl festgelegt. Durch Recherche auf den Internetseiten dieser Hochschulen wurden insgesamt N=847 Kontaktadressen von (1) Hochschulmitarbeiter/-innen mit Aufgaben in der Lehre, (2) Evaluationsprojekten, Stabsstellen zur Durchführung von Evaluation sowie vergleichbaren Strukturen, (3) Fachschaften, die nach eigener Darstellung im Internet Aktivitäten im Bereich der Lehrevaluation unterstützen sowie (4) Sekretariaten ermittelt. Im Zeitraum zwischen Mai und Juni 2006 wurden alle Personen einmalig per E-Mail zur Teilnahme an der Studie eingeladen bzw. um Weiterleitung der Einladung an Lehrende im Arbeitsbereich des jeweiligen Sekretariats gebeten. Auf die Ziele der Befragung wurde hingewiesen; die Teilnahme erfolgte freiwillig und anonym. Insgesamt N=247 Personen (entsprechend einer Teilnahmequote von 29,4%) konnten eindeutig einer der drei Untersuchungsgruppen zugeordnet werden: Als „teilnehmende Studierende“ wurden alle Personen klassifiziert, die in den vergangenen zwei Jahren mehrfach LVE-Fragebögen ausgefüllt hatten, ohne an der LVE-Durchführung (z.B. als Fachschaftsmitglied) beteiligt

gewesen zu sein (n=81, entsprechend 30,7% der ausgewerteten Stichprobe). Der Gruppe „teilnehmende Lehrende“ wurden diejenigen zugeordnet, deren Lehrveranstaltungen in den vergangenen zwei Jahren mehrfach mittels LVE bewertet worden waren, ohne dass die Person selbst an der Auswertung oder Gestaltung der Durchführung beteiligt gewesen war (n=97; 36,7%). Weitere n=69 Personen (26,1%) wurden als „LVE-Anbieter“ eingeordnet, da sie in den vergangenen zwei Jahren an der Gestaltung und/oder Durchführung von Lehrveranstaltungsevaluation beteiligt gewesen waren, zum Beispiel als Fachschaftsmitglieder, Mitarbeiter/-innen von Evaluationsprojekten, LVE-Experten/innen, Evaluationsbeauftragte oder zuständige Mitglieder der Fachbereichs-/Hochschulleitung. Alle anderen abgeschlossenen Fragebögen blieben unberücksichtigt.

Das Durchschnittsalter der Befragten betrug $M=31,4$ Jahre ($SD=9,31$). 26,6% ordneten sich den Naturwissenschaften und Mathematik/Informatik zu, 23,4% entfielen auf Sozialwissenschaften, 10,6% auf Architektur, Ingenieur- sowie Agrarwissenschaften. 7,8% zählten zur Medizin, 7,1% zu den Geisteswissenschaften, 5,7% zu den Rechts- und Wirtschaftswissenschaften, und 7,4% entfielen auf andere Bereiche (n=32 ohne Angabe). 26,6% der Befragten bezogen sich in ihren Einschätzungen auf einen ihnen namentlich nicht weiter bekannten schriftlichen Fragebogen. 20,1% entfielen auf LVE mit einem der publizierten Instrumente (z.B. HILVE, vgl. Rindermann 1997), 12,6% auf dozenten- und 9,6% auf hochschulweit eingesetzte Instrumente. Weitere 8,6% der Befragten beschrieben den Einsatz von fachschaftlich entwickelten Fragebögen, und 7,9% der Befragten bezogen sich auf eine Online-Variante der LVE. 15,1% der Einschätzungen bezogen sich auf andere LVE-Verfahren. Sowohl hinsichtlich der berücksichtigten Fachdisziplinen mit ihren potenziell unterschiedlichen „Evaluationskulturen“ als auch hinsichtlich der einzelnen, den individuellen Erfahrungen zugrunde liegenden LVE-Instrumente wurde somit eine breite Streuung erreicht.

4. Ergebnisse

4.1 Durchführungsmerkmale und Wirkungen von LVE (Fragestellung 1)

Im Fragebogen wurden insgesamt neun Aussagen zur Beschreibung der LVE-Durchführung sowie drei Aussagen zur Einschätzung der Wirkungen vorgegeben. Die Befragten wurden gebeten, das ihnen am besten vertraute LVE-Verfahren mit Hilfe dieser Aussagen auf einer 5-stufigen Skala (1: „stimme gar nicht zu“ bis 5: „stimme völlig zu“) zu beschreiben. Tabelle 1 zeigt die deskriptiven Ergebnisse auf Grundlage aller Befragten (Gesamtstichprobe ohne Differenzierung der Untersuchungsgruppen).

Drei der Durchführungsmerkmale der LVE (Fragestellung 1a) treffen nach Wahrnehmung der Befragten überwiegend zu (Mittelwerte oberhalb des Skalenmittelpunkts, entsprechend $M>3,0$). Demnach beteiligen sich die Studierenden zumeist gerne an der LVE (Aussage 1, $M=3,59$), sie nutzen die LVE in der Regel nicht für persönliche, nicht-konstruktive Kritik (Aussage 3, umgepoltes Item: höhere Werte stehen für stärkere Zurückweisung der Aussage, $M=3,59$), und die Dozenten/innen setzen sich für die meisten erkennbar

Tabelle 1: Beobachtete Durchführungsmerkmale und wahrgenommene Wirkungen der Lehrveranstaltungsevaluation in der Gesamtstichprobe (1: „stimme gar nicht zu“ bis 5: „stimme völlig zu“; Prozentbasis N umfasst jeweils abgegebene Einschätzungen ohne „keine Angabe“)

Eingeschätzter Aspekt	N	M	SD	Verteilung der Antwortkategorien				
				%1	%2	%3	%4	%5
<i>Durchführungsmerkmale</i>								
1 Die meisten Studierenden beteiligen sich gerne an der Lehrveranstaltungsevaluation.	234	3,59	1,02	3,4	12,4	22,6	44,9	16,7
2 Die meisten Lehrenden beteiligen sich gerne an der Lehrveranstaltungsevaluation.	205	3,00	1,01	6,8	23,4	38,5	24,9	6,3
3 Viele Studierende nutzen die LVE für persönliche, nicht-konstruktive Kritik. (umgepoltes Item)	229	3,59	1,07	4,4	10,0	29,7	34,1	21,8
4 Die Dozenten/innen setzen sich erkennbar mit Evaluationsergebnissen auseinander.	223	3,19	1,05	5,4	22,4	28,7	34,5	9,0
5 Die Lehrenden stellen in der Regel die Evaluationsergebnisse den Studierenden persönlich vor.	220	2,55	1,37	31,4	20,5	21,4	15,5	11,4
6 Die Evaluationsergebnisse werden (durch Aushang, Internet) im Regelfall veröffentlicht.	217	2,61	1,55	35,0	22,6	8,8	13,8	19,8
7 Aus den Ergebnissen wird eine Rangfolge der Lehrenden ermittelt.	191	1,91	1,42	63,4	12,0	6,8	5,8	12,0
8 Im Anschluss an LVE stehen Maßnahmen (Beratung, Workshops) für die Lehrenden bereit.	181	1,78	1,28	64,1	16,0	5,0	7,2	7,7
9 LVE ist eingebettet in ein Rahmenkonzept zur Entwicklung der (didaktischen) Kompetenzen der Lehrenden.	181	2,03	1,26	48,6	22,7	11,0	12,2	5,5
<i>Wirkungsdimensionen</i>								
10 Insgesamt gesehen profitieren die Studierenden von der Durchführung.	232	3,34	1,16	7,8	18,1	21,1	38,8	14,2
11 Insgesamt gesehen profitieren die Lehrenden von der Durchführung.	219	3,62	1,07	4,1	12,3	21,9	41,1	20,5
12 Durch die LVE kommt es zu einer spürbaren Verbesserung der Veranstaltungen.	217	2,66	1,11	17,5	27,6	30,0	21,2	3,7

mit Evaluationsergebnissen auseinander (Aussage 4, M=3,19). Drei weitere Beschreibungen treffen mit durchschnittlichen Einschätzungen in der Nähe und knapp unterhalb des Skalenmittelpunkts zumindest teilweise zu: Viele Lehrende nehmen gerne an der LVE teil (Aussage 2, M=3,00), eine Reihe von ihnen stellt die Evaluationsergebnisse den Studierenden persönlich vor (Aussage 5, M=2,55), und in vielen Fällen werden die Evaluationsergebnisse veröffentlicht (Aussage 6, M=2,61). Die verbleibenden drei Durchführungsmerkmale können von vielen Befragten überhaupt nicht eingeschätzt werden (hoher Anteil „keine Angabe“) und werden von den anderen überwiegend zurückgewiesen (Mittelwerte um und unterhalb des ersten ablehnenden Skalenpunkts, entsprechend 2,0): Aus den Ergebnissen wird zumeist keine Rangfolge der Lehrenden ermittelt (Aussage 7, M=1,91), Anschlussmaßnahmen an LVE stehen in der Regel nicht bereit (Aussage 8, M=1,78), und LVE ist überwiegend nicht eingebettet in ein Rahmenkonzept zur didaktischen Kompetenzentwicklung (Aussage 9, M=2,03). Auch hinsichtlich der Wirkungen (Fragestellung 1b) wird von den Befragten eine zweiseitige Wahrnehmung angegeben: Studierende (Aussage 10, M=3,34) und, noch ausgeprägter, Lehrende (Aussage 11, M=3,62) profitieren zumindest überwiegend von der LVE-Durchführung. Demgegenüber wird die Verbesserung der Lehrveranstaltungen (Aussage 12, M=2,66) als wenig spürbar eingestuft; die durchschnittliche Einschätzung liegt hier klar im ablehnenden Abschnitt der Antwortskala.

4.2 Wahrnehmungsunterschiede zwischen Studierenden, Lehrenden und Evaluationsanbietern (Fragestellung 2)

Um etwaige Unterschiede zwischen den drei Perspektiven zu ermitteln, wurden die Einschätzungen zu den Durch-

führungsmerkmalen und den Wirkungsdimensionen jeweils als abhängige Variablen in eine univariate Varianzanalyse (ANOVA) mit dem Faktor Perspektive (3-stufig: Studierende, Lehrende sowie Evaluationsanbieter) eingebracht. Im Falle eines (ggf. annähernd) signifikanten Haupteffekts wurden post hoc-Tukey-Tests vorgenommen, um die einzelnen Untersuchungsgruppen miteinander zu vergleichen. Tabelle 2 zeigt die Ergebnisse dieser Analysen.

Hinsichtlich der LVE-Durchführungsmerkmale unterscheiden sich die untersuchten Gruppen in vier der neun Aspekte signifikant; bei zwei weiteren Aussagen sind im Vergleich der Mittelwerte statistisch nicht-signifikante Trends zu beobachten. Am ehesten sind es die Studierenden selbst, die eine Nutzung der LVE für persönliche, nicht-konstruktive Kritik beobachten (Aussage 3, umgepoltes Item: höhere Werte entsprechen stärkerer Zurückweisung der Aussage). Besonders die Lehrenden geben an, sich erkennbar mit den Evaluationsergebnissen auseinanderzusetzen (Aussage 4) und die Ergebnisse persönlich mit den Studierenden zu diskutieren (Aussage 5). LVE-Anbieter geben hingegen verstärkt an, dass Ergebnisse veröffentlicht werden (Aussage 6), dass eine Rangfolge der Lehrenden ermittelt wird (Aussage 7) und dass entsprechende LVE-Anschlussmaßnahmen bereitgestellt werden (Aussage 8).

Hinsichtlich der Wirkungsdimensionen ergeben sich bei allen drei eingeschätzten Aspekten signifikante Unterschiede zwischen den drei Gruppen: Die Bewertung aus Sicht der Studierenden liegt jeweils in einem kritischeren Bereich als die entsprechenden Einschätzungen der Lehrenden und der LVE-Anbieter. Dies gilt sowohl für die Frage, ob Studierende von der LVE-Durchführung profitieren (Aussage 10, mit M=2,96 nur teilweise Bestätigung aus Sicht der Studierenden), als auch für den wahrgenommenen Nutzen der LVE für die Lehrenden (Aussage 11) und ebenso für die etwaige Spürbarkeit der Verbesserung von Lehrveranstaltungen durch LVE (Aussage 12, mit M=2,15 klare Ablehnung dieser Aussage durch die Studierenden).

4.3 Unterschiedliche Erklärungsmuster der Beteiligten (Fragestellung 3)

Abschließend sollte nach Hinweisen dafür gesucht werden, dass die unterschiedlichen Wahrnehmungen der „Stakeholder“ auch von unterschiedlichen Erklärungsmustern über Lehrveranstaltungsevaluation geprägt werden. Im Fokus dieser Erklärungsmuster steht der vom Individuum wahrgenommene Zusammenhang zwischen den Durchführungs-

Tabelle 2: Vergleich der drei Perspektiven (N=247, Fälle mit fehlenden Werten ausgeschlossen; Signifikanzniveaus: n.s. nicht signifikant: $p > 0,10$; (*) marginal signifikant: $p < 0,10$; *signifikant: $p < 0,05$; (<) post hoc-Vergleich marginal signifikant: $p < 0,10$; < post hoc-Vergleich signifikant: $p < 0,05$)

Eingeschätzter Aspekt	Studierende (n=81)		Lehrende (n=97)		Anbieter (n=69)		Vergleich		
	M	SD	M	SD	M	SD	F	p	post-hoc
<i>Durchführungsmerkmale</i>									
1 Teilnahmemotivation der Studierenden	3,41	1,07	3,69	0,98	3,67	0,98	1,85	0,16 n.s.	-
2 Teilnahmemotivation der Lehrenden	2,86	0,90	3,05	0,99	3,07	1,12	0,83	0,437 n.s.	-
3 Nutzung für persönliche, nicht-konstruktive Kritik (U)	3,28	1,04	3,78	1,05	3,67	1,07	4,80	0,009 *	S<L, S(<)A
4 Erkennbare Auseinandersetzung der Lehrenden	2,96	1,18	3,41	0,99	3,17	0,94	3,72	0,026 *	S<L
5 Persönliche Vorstellung der Ergebnisse	2,32	1,35	2,80	1,39	2,51	1,32	2,51	0,084 (*)	S(<)L
6 Veröffentlichung der Ergebnisse	2,55	1,43	2,36	1,55	3,00	1,64	3,16	0,044 *	L<A
7 Ermittlung einer Rangfolge der Lehrenden	1,88	1,41	1,70	1,33	2,22	1,52	2,32	0,101 (*)	L(<)A
8 Bereitstellung von LVE-Anschlussmaßnahmen	1,58	1,05	1,57	1,10	2,20	1,53	4,99	0,008 *	S(<)A, L<A
9 Einbettung der LVE in ein Rahmenkonzept	1,92	0,97	2,04	1,32	2,10	1,34	0,23	0,792 n.s.	-
<i>Wirkungsdimensionen</i>									
10 Nutzen für Studierende	2,96	1,19	3,41	1,12	3,67	1,06	7,42	0,001 *	S<L, S<A
11 Nutzen für Lehrende	3,33	1,14	3,72	1,02	3,75	1,03	3,17	0,044 *	S(<)L, S(<)A
12 Verbesserung von Veranstaltungen	2,15	0,99	2,88	1,09	2,93	1,07	12,10	0,001 *	S<L, S<A

merkmalen der LVE einerseits und ihren Wirkungsdimensionen andererseits. Voraussetzung für die weiterführende Untersuchung dieser Erklärungsmuster ist das Vorliegen von Unterschieden in den Einschätzungen zwischen den drei Perspektiven, die über etwaige Niveauunterschiede hinausgehen. Um diese Voraussetzung zu prüfen, wurde zunächst eine 2-faktorielle Varianzanalyse mit dem Zwischensubjektfaktor Perspektive (3-stufig: Studierende, Lehrende, LVE-Anbieter) sowie dem Innersubjektfaktor Merkmalsdimension (12-stufig: 9 Durchführungs- sowie 3 Wirkungsdimensionen) und der individuellen Einschätzung des jeweiligen Merkmals auf der 5-stufigen Antwortskala als abhängiger Variable berechnet. Eine signifikante Interaktion von Perspektive und Merkmalsdimension wäre ein Hinweis auf unterschiedliche Wahrnehmungsmuster und damit die Grundlage für weiterführende Analysen. Einzelne fehlende Werte wurden jeweils durch den Gesamtmittelwert entsprechend Tabelle 1 ersetzt. Die Varianzanalyse erbrachte zwei signifi-

kannte Haupteffekte, nämlich für die Perspektive ($F(2, 244)=4,61$; $p=0,011$) und für die Merkmalsdimension ($F(11, 2684)= 86,02$; $p < 0,001$). Die entscheidende Interaktion erwies sich ebenfalls als signifikant ($F(22, 2684)=3,67$; $p < 0,001$), sodass die Annahme von Unterschieden, die über Niveauunterschiede hinausgehen, aufrecht erhalten werden kann. Ausgehend von diesem Befund, wurden zur Beschreibung der Erklärungsmuster die Zusammenhänge zwischen den Durchführungsmerkmalen und den Wirkungsdi-

Tabelle 3: Regressionsmodelle zur Abbildung der Erklärungsmuster (abschließende Modelle nach Aufnahme aller signifikanten Prädiktoren; 1 Teilnahmemotivation der Studierenden, 2 Teilnahmemotivation der Lehrenden, 3 Nutzung für nicht-konstruktive Kritik, 4 Auseinandersetzung der Lehrenden mit Ergebnissen, 5 persönliche Ergebnisvorstellung durch Lehrende, 6 Veröffentlichung der Ergebnisse, 7 Rangfolge der Lehrenden, 8 Bereitstellung von Anschlussmaßnahmen, 9 Einbettung in Rahmenkonzept; alle Modelle *signifikant mit $p < 0,01$)

Kriterien (Wirkungsdimensionen)	beta-Gewichte der aufgenommenen Prädiktoren (Durchführungsmerkmale)									Modellprüfung		
	1	2	3	4	5	6	7	8	9	F	df	R ²
<i>nur Studierende (n=81)</i>												
Nutzen für Studierende				0,45					0,30	21,81	2,78	0,36*
Nutzen für Lehrende		0,22		0,47					0,21	20,48	3,77	0,44*
Verbesserung von Veranstaltungen	0,21			0,61				0,16		33,50	3,77	0,57*
<i>nur Lehrende (n=97)</i>												
Nutzen für Studierende	0,29				0,32				0,30	9,01	3,93	0,23*
Nutzen für Lehrende	0,42			0,21	0,22		-0,24		0,32	13,10	5,91	0,42*
Verbesserung von Veranstaltungen	0,43			0,19	0,21				0,28	11,97	4,92	0,34*
<i>nur LVE-Anbieter (n=69)</i>												
Nutzen für Studierende		0,30		0,30	0,29					16,92	3,65	0,44*
Nutzen für Lehrende		0,38							0,45	18,50	2,66	0,36*
Verbesserung von Veranstaltungen		0,42							0,33	13,73	2,66	0,29*
<i>Gesamtstichprobe (N=247)</i>												
Nutzen für Studierende	0,20			0,30	0,14				0,21	23,99	4,242	0,28*
Nutzen für Lehrende	0,19	0,12	-0,10	0,25	0,13				0,29	15,97	6,240	0,36*
Verbesserung von Veranstaltungen	0,24			0,40		0,15			0,21	28,46	4,242	0,37*

mensionen innerhalb der drei Untersuchungsgruppen jeweils separat weiter analysiert. Für die Daten aus jeder der drei Gruppen einzeln sowie, als ergänzende Information, auf Basis der Daten aller Befragten wurden schrittweise lineare Regressionen berechnet, bei denen jeweils eine der drei Wirkungsdimensionen durch einzelne oder mehrere Durchführungsmerkmale vorhergesagt werden sollte. Wiederum wurden einzelne fehlende Werte

durch den Gesamtmittelwert entsprechend Tabelle 1 ersetzt; die Ergebnisse zeigt Tabelle 3.

Bei den drei Modellen, welche die Erklärungsmuster in der Gruppe der Studierenden abbilden, erwies sich durchweg die Auseinandersetzung der Dozenten/innen mit den Ergebnissen (Durchführungsmerkmal 4, $\beta=0,45$, $0,47$ bzw. $0,61$) als signifikanter und jeweils stärkster Prädiktor für die Ergebnismerkmale. Zusätzlich wurde die Einbettung der LVE in ein Rahmenkonzept zur didaktischen Kompetenzentwicklung der Lehrenden (Durchführungsmerkmal 9, $\beta=0,30$ bzw. $0,21$) in zwei der Modelle als signifikanter Prädiktor identifiziert, nicht jedoch bei der Frage, ob sich die Veranstaltungen spürbar verbessern. Bei der Modellierung der Lehrendenperspektive zeigten sich hingegen die Teilnahmemotivation der Studierenden (Durchführungsmerkmal 1, $\beta=0,29$, $0,42$ bzw. $0,43$), die persönliche Vorstellung der Ergebnisse gegenüber den Teilnehmenden (Durchführungsmerkmal 5, $\beta=0,32$, $0,22$ bzw. $0,21$)

sowie die Einbettung in ein Rahmenkonzept (Durchführungsmerkmal 9, $\beta=0,30$, $0,32$ bzw. $0,28$) durchweg als signifikante Prädiktoren zur Erklärung der Wirkungsdimensionen. Beim Nutzen für die Lehrenden wurde zudem die eigene Auseinandersetzung mit den Ergebnissen (Merkmal 4, $\beta=0,21$) sowie – als Prädiktor mit negativem Einfluss auf die Wirkungsdimension – die Veröffentlichung von Rankings (Merkmal 7, $\beta=-0,24$) als signifikant herausgestellt. Aus den Einschätzungen der LVE-Anbieter erweist sich hingegen durchweg die Teilnahmemotivation der Lehrenden (Durchführungsmerkmal 2; $\beta=0,30$, $0,38$ bzw. $0,42$) als einflussreich für die Wirkung der LVE. In zwei der drei Modelle ist darüber hinaus die Einbettung in ein Rahmenkonzept (Merkmal 9; $\beta=0,45$ bzw. $0,33$) relevant, und hinsichtlich des Nutzens für die Studierenden erwiesen sich aus Sicht der LVE-Anbieter die Auseinandersetzung der Lehrenden mit den Ergebnissen (Merkmal 4, $\beta=0,30$) sowie deren persönliche Ergebnisvorstellung (Merkmal 5, $\beta=0,29$) als weitere relevante Einflussfaktoren. Alle Regressionsmodelle weisen mit $R^2=0,23\dots0,57$ eine substantielle Erklärungskraft auf und bieten jeweils einen statistisch signifikanten Aufklärungsbeitrag.

4.4 Zusammenfassung der Ergebnisse

Die Befragten zeichnen ein gemischtes Bild vom aktuellen Stand der Lehrveranstaltungsevaluation an den deutschen Hochschulen. Dies gilt zunächst hinsichtlich der Durchführungsmerkmale (Fragestellung 1a, vgl. Tabelle 1): Studierende beteiligen sich überwiegend, aber nicht durchweg, gerne an der LVE und vermeiden es überwiegend, aber nicht durchweg, die Gelegenheit zu einer persönlichen, nicht-konstruktiven Kritik zu nutzen. Dozentinnen und Dozenten setzen sich eher erkennbar als nicht erkennbar mit den Ergebnissen auseinander. Rangfolgen der Lehrenden werden in der Regel nicht erstellt, und eine Einbettung in ein Rahmenkonzept zur (hochschuldidaktischen) Kompetenzentwicklung ist für die meisten ebenso wenig erkennbar wie im Anschluss an LVE bereitstehende Trainings- oder Entwicklungsmaßnahmen. Entsprechend werden auch die Wirkungsdimensionen der Lehrveranstaltungsevaluation eher verhalten bewertet (Fragestellung 1b): Während sowohl die Lehrenden als auch die Studierenden noch überwiegend, wenn auch nicht uneingeschränkt, von der Durchführung von LVE profitieren, bleiben konkrete Verbesserungen von Lehrveranstaltungen aufgrund von LVE oftmals aus – sie sind im Durchschnitt eher weniger als stärker spürbar, und für 17,5% der Befragten überhaupt nicht (vgl. Tabelle 1).

Allerdings gibt es deutliche Unterschiede in den Einschätzungen sowohl der einzelnen Verfahren (vgl. Häufigkeitsverteilung in Tabelle 1) als auch zwischen den hier betrachteten drei Perspektiven (Fragestellung 2).

So schätzen die Studierenden die drei Wirkungsdimensionen (vgl. Tabelle 2) durchweg kritischer ein als die Lehrenden bzw. die LVE-Anbieter. Hinsichtlich der Durchführungsmerkmale ergibt sich ein noch weiter differenziertes Bild: Lehrende empfinden deutlich stärker als die anderen beiden Gruppen, dass sie sich mit den Evaluationsergebnissen „erkennbar“ auseinandersetzen und beobachten deutlicher als Studierende und LVE-Anbieter, dass sie und ihre Kolleginnen und Kollegen die Ergebnisse persönlich

den Teilnehmenden vorstellen. Studierende haben stärker als die anderen beiden Gruppen den Eindruck, dass LVE für persönliche, nicht-konstruktive Kritik an den Lehrenden genutzt wird. LVE-Anbieter schließlich nehmen mehr als die anderen befragten Gruppen wahr, dass die Evaluationsergebnisse veröffentlicht werden, dass Rangfolgen auf Basis der Studierendeneinschätzungen erstellt werden und dass Anschlussmaßnahmen wie Workshops oder Seminare zur Verfügung stehen.

Diese Unterschiedlichkeit in der Einschätzung der Durchführungs- und Ergebnismerkmale setzt sich in der Analyse der Erklärungsmuster (Fragestellung 3, Tabelle 3) konsequent fort: Studierende beobachten vor allem dann eine erfolgreiche Lehrveranstaltungsevaluation, wenn sich die Dozentinnen und Dozenten erkennbar mit den LVE-Ergebnissen auseinandersetzen. Lehrende hingegen nehmen LVE vor allem dann als wirksam wahr, wenn auch die Studierenden sich gerne an der Evaluation beteiligen und wenn sie selbst – die Lehrenden – ihre Ergebnisse den Teilnehmenden persönlich vorstellen. LVE-Anbieter hingegen sehen vor allem einen Zusammenhang zwischen der Motivation der Lehrenden, sich an LVE zu beteiligen, und deren Wirksamkeit. Doch trotz dieser Unterschiedlichkeit gibt es in einem wesentlichen Punkt auch eine Übereinstimmung: Alle drei Wirkungsdimensionen werden maßgeblich davon beeinflusst, inwieweit die jeweils durchgeführte Variante der Lehrveranstaltungsevaluation in ein Rahmenkonzept zur lehrbezogenen Kompetenzentwicklung der Dozentinnen und Dozenten eingebettet ist, was gemäß der vorliegenden Beobachtungen (vgl. Tabelle 1) lediglich in 5,5% der Fälle vollständig erfüllt ist – in 48,6% hingegen überhaupt nicht.

Einschränkend ist anzumerken, dass die Stichprobe kein repräsentatives Bild „der“ Lehrveranstaltungsevaluation an deutschen Hochschulen abbilden kann. Zum einen ist dies bedingt durch die Selbstselektion der Stichprobe, indem nur diejenigen an der Befragung teilgenommen haben, die ein Interesse daran hatten, ihre Einschätzungen über LVE abzugeben. Mit den Nicht-Nutzern (d.h. denjenigen Studierenden und Lehrenden, die sich aus bestimmten, individuell unterschiedlichen Gründen nicht beteiligen) und den Nicht-Anbietern (d.h. denjenigen Personen, Projekten und Einrichtungen, die LVE anbieten und durchführen könnten, dies aber nicht tun) sind ganz wesentliche, mutmaßlich kritischere Gruppen von der Studie ausgeschlossen worden, indem ausschließlich die Daten derjenigen berücksichtigt wurden, die innerhalb der zurückliegenden zwei Jahre persönlich an LVE beteiligt waren.

Auch wurde keine exakte Abbildung der an deutschen Hochschulen tatsächlich eingesetzten LVE-Verfahren in der Studie realisiert, d.h. die Aussagen beziehen sich auf eine große Bandbreite unterschiedlicher Instrumente und Durchführungsvarianten, die nicht repräsentativ für die „durchschnittliche“ Form der Lehrveranstaltungsevaluation an deutschen Hochschulen sein muss.

Dennoch zeigen sich statistisch signifikante und inhaltlich bedeutsame Unterschiede in den Bewertungen, die sich nicht ausschließlich durch eine etwaige selektive Verzerrung (z.B. generell kritischere Einstellung der Studierenden, vgl. Tabelle 2) erklären lassen.

5. Diskussion

Selbst über scheinbar einfache und objektive Fakten herrscht keine Einigkeit zwischen Studierenden, Lehrenden und den Anbietern von Lehrveranstaltungsevaluation: Werden die LVE-Ergebnisse nun veröffentlicht, oder werden sie es nicht? Setzen sich Lehrende erkennbar mit den Evaluationsergebnissen auseinander, oder erfolgt diese Auseinandersetzung eher unsichtbar, wenn überhaupt? Stehen Anschlussmaßnahmen tatsächlich bereit, oder vertrauen die LVE-Anbieter lediglich darauf, dass es solche Maßnahmen an der jeweiligen Hochschule gibt? Diese Wahrnehmungen scheinen auch von der jeweiligen Perspektive beeinflusst zu werden: Niemand ist stärker als die Studierenden selbst bereit, an der Konstruktivität des studentischen Feedbacks zu zweifeln, niemand sieht die Lehrenden so engagiert und auseinandersetzungsbereit wie diese sich selbst, und niemand nimmt die Ausschöpfung der Auswertungspotenziale und das Angebot an Entwicklungsangeboten so umfassend wahr wie die LVE-Anbieter selbst. Sogar die Erklärungsmuster für wirkungsvolle Lehrveranstaltungsevaluation unterscheiden sich in einer vor dem Hintergrund der subjektiven und objektiven Realitäten (vgl. Rindermann 2000) verständlichen Weise. Nur in einem Punkt herrscht weitgehende und ernüchternde Einigkeit: Aufgrund der nur punktuell gegebenen Einbettung der Lehrveranstaltungsevaluation in ein für die Beteiligten erkennbares, nachvollziehbares, akzeptables Rahmenkonzept fallen der Nutzen für die Studierenden und die Lehrenden sowie ganz besonders die aufgrund von LVE spürbaren Verbesserungen von Lehrveranstaltungen bislang eher moderat als triumphal aus.

Was folgt aus diesen Erkenntnissen? Drei Empfehlungen lassen sich ableiten: (1) Bei der Lehrveranstaltungsevaluation geht es zentral um den Austausch von Perspektiven, um die gemeinsame Reflexion der in der Hochschullehre ablaufenden Prozesse, mit dem Ziel, Erwartungen, objektive und subjektive Realität immer wieder aufeinander zu beziehen. Offenbar ist noch mehr Transparenz auch über den Prozess der Lehrveranstaltungsevaluation selbst herstellbar. Wozu soll die Befragung der Studierenden dienen, woran wird erkennbar, wenn Lehrende sich mit den Ergebnissen auseinandersetzen, welche Anschlussmaßnahmen soll es konkret geben? Die drei in dieser Studie betrachteten Perspektiven unterscheiden sich in ihren Schlussfolgerungen und auch in den Erklärungsmustern, die zwischen Durchführung und Wirkung der LVE gesehen werden. Eine große Chance kann darin liegen, eine Verständigung auch auf dieser Ebene anzustreben, um ein stärker als bislang geteiltes Verständnis der Ziele, des Nutzens und der Wirkungsweise von LVE zu erreichen. (2) Als „Anwälte der Lehrveranstaltungsevaluation“ können die Anbieter – von fachschaftlichen Evaluationsprojekten mit einem studierendenzentrierten Fokus über Vorhaben, die auf der Ebene von Fachbereichen oder Fakultäten angesiedelt sind, bis hin zu hochschulweiten oder gar hochschulübergreifenden Aktivitäten – in besonderem Maße Einfluss auf die Durchführung von LVE nehmen. Folgen sie ausschließlich ihrem eigenen Erklärungsmodell, in dem die Teilnahmemotivation der Lehrenden den umfassendsten Zusammenhang mit den Wirkungsdimensionen aufweist, gehen sie damit das Risiko ein, die von den anderen Perspektiven als wesentlich erachteten

Durchführungsmerkmale zu vernachlässigen. Wenn es im Interesse der LVE-Anbieter liegt, die Wahrnehmung über LVE und die auch für die anderen Beteiligten erkennbare Wirksamkeit positiv zu steuern, sind auch die von den Studierenden und den Lehrenden als ausschlaggebend angesehenen Punkte aufzugreifen, um den LVE-Prozess nicht nur wirksam, sondern auch plausibel zu machen. Das kann geschehen durch einfache, aber als effektiv erlebte Fragebogengestaltung, durch die Bereitstellung von Präsentationsvorlagen für eine schnelle und unkomplizierte Ergebnisdiskussion im Hörsaal oder durch Maßnahmen, die das Image der Lehrveranstaltungsevaluation positiv beeinflussen.

(3) Fast trivial und doch als bedeutsam zeigt sich erneut die Relevanz eines Rahmenkonzepts zur (hochschuldidaktischen) Kompetenzentwicklung von Lehrenden. Dort, wo es solche Konzepte gibt, wird LVE offenbar als effektiv wahrgenommen – kein anderer Einzelprädiktor ist insgesamt und durch die Erklärungsmuster aller Perspektiven hindurch so relevant für die Wirkung von Lehrveranstaltungsevaluation. Im konstruktiven Austausch der „Stakeholder“, hier am Beispiel der Studierenden, Lehrenden und LVE-Anbieter untersucht, aber auch darüber hinaus: der Hochschulleitungen, der didaktischen Einrichtungen, externen Geldgeber für entsprechende Modellprojekte, sind daher solche Konzepte zu entwickeln, die nicht nur von einzelnen, sondern von den meisten als plausibel und gewinnbringend erlebt werden können. Mehrere Erfolgsfaktoren ergeben sich aus der vorliegenden Studie: Ein Rahmenkonzept kann die Wirksamkeit von LVE dann steigern, wenn es die Sichtbarkeit der Auseinandersetzung der Lehrenden mit den Ergebnissen erhöht, die Konstruktivität des Feedbacks weiter unterstützt, den Dialog mit den Studierenden fördert und nachvollziehbare Verknüpfungen zwischen LVE-Ergebnissen und Entwicklungsangeboten schafft. An erster Stelle steht hierbei die Akzeptanz durch jeweils diejenigen, die maßgeblich beteiligt sind: Wenn Studierende ihren Beitrag zur LVE – das Ausfüllen von Fragebögen, ggf. die Beteiligung an einer Ergebnisdiskussion – als sinnvoll erleben, wenn Lehrende keine Sorge vor allzu öffentlichen Rückmeldungen haben und auf einfache Weise in einen konstruktiven, ergebnisorientierten Dialog mit ihren Studierenden treten können und wenn LVE-Anbieter die Gelegenheit erhalten, für den Sinn und Nutzen von Lehrveranstaltungsevaluation zu werben und in eine Zusammenarbeit mit anderen Stellen der jeweiligen Hochschule (z.B. Nachwuchsförderung, Qualitätsentwicklung, Hochschuldidaktik, vgl. Schmidt 2007) zu treten, um gemeinsame Elemente eines Rahmenkonzepts zur Kompetenzentwicklung von Lehrenden aufzubauen – wenn diese Bedingungen schrittweise in den Fokus genommen werden, kann Lehrveranstaltungsevaluation stärker als bislang ihre Potenziale entfalten und zu einer bisher nur bruchstückhaften Verständigung zwischen den unterschiedlichen Perspektiven beitragen.

Literaturverzeichnis

- Aleamoni, L.M. (1987):* Typical faculty concerns about student evaluation of teaching. *New Directions for Teaching and Learning*, 31, pp. 25-31.
- Benz, W./Kohler, J./Landfried, K. (Hg.) (2004):* Handbuch Qualität in Studium und Lehre: Evaluation nutzen - Akkreditierung sichern - Profilschärfen. Berlin.
- Cohen, P. A. (1980):* Effectiveness of student-rating feedback for improving college instruction: A metaanalysis of findings. *Research in Higher Education*, 13, pp. 321-341.

- Daxner, M. (1999):* Evaluation, Indikatoren und Akkreditierung: Auf dem Weg in die Rechtfertigungsgesellschaft. In HRK (Hg.): Viel Lärm um nichts? Evaluation von Studium und Lehre und ihre Folgen. Beiträge zur Hochschulpolitik 4/1999, Bonn, S. 41-49.
- ESIB (Hg.) (2003):* Student Bologna Surveys. Brüssel: ESIB. The National Unions of Students in Europe.
- Hage, N. el (1996):* Lehrevaluation und studentische Veranstaltungskritik: Projekte, Instrumente und Grundlagen. Bonn.
- Marsh, H. W. (1984):* Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, Vol. 75/No. 5, pp. 707-754.
- Rindermann, H. (1997):* Die studentische Beurteilung von Lehrveranstaltungen: Forschungsstand und Implikationen für den Einsatz von Lehrevaluationen. In: Jäger R./Trost, G./Lehmann, R. (Hg.): Tests und Trends 11: Jahrbuch der Pädagogischen Diagnostik, Weinheim, S. 12-53.
- Rindermann, H. (1999):* Was zeichnet gute Lehre aus? Ergebnisse einer offenen Befragung von Studierenden und Lehrenden. *Zeitschrift für Hochschuldidaktik*, Jg. 23/H. 1, S. 136-156.
- Rindermann, H. (2000):* Das Selbstobjektivierungsproblem im akademischen Milieu (I). In: *Das Hochschulwesen*, Jg. 48/H. 3, S. 74-82.
- Schmidt, B. (2007):* Die strukturelle Basis der Hochschulentwicklung an deutschen Hochschulen – eine quantitative Internetrecherche. In: *Hochschulmanagement*, Jg. 2/H. 1, S. 22-29.
- Schmidt, B./Loßnitzer, T. (2007):* Konzepte für ein fachinternes Qualitätsmanagement der Lehre – Das Modell des Universitätsprojekts Lehrevaluation an der Friedrich-Schiller-Universität Jena. In: *Qualität in der Wissenschaft*, Jg. 1/H. 2, S. 44-50.
- Schulze, T. (2001):* Was erwarten die Studierenden von der Evaluation? In: HRK (Hg.): Auf dem Weg zum Qualitätsmanagement: Erfahrungen und Perspektiven, Beiträge zur Hochschulpolitik. Bonn, S. 17-18.
- Schweer, M./Rosemann, B. (1995):* Qualität der Lehre: Bedingungsvariablen des studentischen Urteils. In: *Zeitschrift für Pädagogische Psychologie*, 9 (3/4), S. 189-196.

■ Dr. Boris Schmidt, Dipl.-Kfm., Dipl.-Psych.,
Universitätsprojekt Lehrevaluation, Friedrich-
Schiller-Universität Jena,
E-Mail: schmidt@thema31.de

Sehr geehrte Leserinnen und Leser,

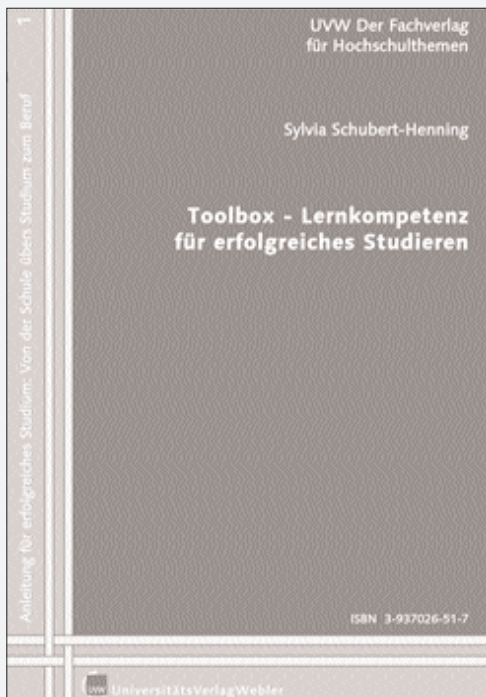
nachfolgend bietet der UniversitätsVerlagWebler einen Titel als ideale Ratgeberliteratur für den Ausbau der Tutorenprogramme an Hochschulen an.

Fachbereiche geben dieses Buch mit Hilfe der Studienbeiträge jedem Tutor als persönliche Ausstattung an die Hand („Verbrauchsmaterial“).

Das ist eine einmalige Investition in die Qualität der Tutorien, die den Studierenden direkt zugute kommt, also sich lohnt!

Sylvia Schubert-Henning

Toolbox - Lernkompetenz für erfolgreiches Studieren



Die „Toolbox – Lernkompetenz für erfolgreiches Studieren“ enthält 40 Tools, die lernstrategisches Know-how für selbstgesteuertes Lernen mit Techniken des wissenschaftlichen Arbeitens verknüpfen. Diese Handwerkszeuge unterstützen Studierende bei der Verbesserung ihres Selbstmanagements, beim gezielten Lesen von wissenschaftlichen Texten sowie beim Vorbereiten und der Präsentation von Referaten. Darüber hinaus erhalten Studierende mit den Tools grundlegende Tipps zum Erstellen von Hausarbeiten oder zur Prüfungsvorbereitung. Die Tools eignen sich besonders gut als kompaktes Material für Fachtutorien in der Studienanfangsphase, für selbstorganisierte Lerngruppen oder auch für Studierende, die sich diese Fertigkeiten im Selbststudium aneignen wollen. Ein Blick auf die theoretischen Grundlagen von Lernkompetenzen lassen die Werkzeuge des selbstgesteuerten Lernens im Studium „begreifbar“ werden. Mit einer gezielten Anwendung der Tools werden die Lernmotivation und die Freude am Studieren maßgeblich gestärkt.

ISBN 3-937026-51-7, Bielefeld 2007, 110 Seiten, 14.60 Euro

Bestellung

Mail: info@universitaetsverlagwebler.de,

Fax: 0521/ 923 610-22

Reihe Anleitung für erfolgreiches Studium:
Von der Schule übers Studium zum Beruf



*Küchler, Uwe (2007):
Interkulturelle Hochschullehre:
Internationalisierung am Beispiel
der Amerikanistik.*

Münster, LIT Verlag.
ISBN: 3-825804-86-0,
336 Seiten, 28.90 Euro

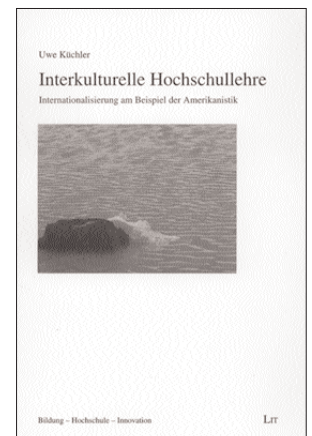
Der neue Blickwinkel der jüngeren Hochschuldidaktik, der eine Verschiebung vom Lehren hin zum Lernen postuliert, wird im vorliegenden Band um eine interkulturelle Komponente erweitert. In Form einer ethnographisch-explorativen Fallstudie im Institut für Amerikanistik in einer deutschen Universität untersucht der Autor das Konzept des intercultural classroom sowie innovative Lehr- und Lernformate.

Uwe Küchler verschränkt theoretische und methodische Ansätze aus der Hochschuldidaktik, aus den American Studies und aus der qualitativen Sozialforschung. Nach einer kurzen Beschreibung der im Rahmen des Bologna-Prozesses angestrebten, aber häufig nicht umgesetzten Internationalisierung deutscher Universitäten zeichnet der Autor ein Bild interkultureller Hochschullehre, indem er an der seit gut einem Jahrzehnt vorgeschlagenen „Orientierung der Lehre am Lernen“ ansetzt. Es ist ihm hoch anzurechnen, dass er die auch im deutschen Diskurs übliche Formulierung *Shift from teaching to learning* auf eine verständliche deutsche Formel gebracht hat, die er auf die speziellen Bedürfnisse der interkulturellen und interdisziplinären Amerikanistik anwendet.

Er verfolgt dabei die Frage, mit welchen hochschuldidaktischen Ansätzen eine interkulturelle Orientierung im Studium sich bestmöglich verwirklichen lässt. Diese untersucht Küchler im Rahmen einer ethnographisch-explorativen Fallstudie unterschiedlicher Lehr-/Lernformate am Institut für Amerikanistik an einer deutschen Universität. Er setzte zur Datenerhebung problemzentrierte Interviews mit Studierenden und Lehrenden ein, flankiert von teilnehmender Beobachtung in den intercultural classrooms.

In der Publikation ist die methodische Anlage der Studie nur kurz dargestellt, so dass einige Untersuchungs- und Auswertungsschritte in der Ergebnisdarstellung unklar bleiben, beispielsweise der Einsatz eines Beobachtungs- und Interviewleitfadens bzw. die analytische Trennung der Interviews mit den Studierenden und Lehrenden.

Diese Intransparenz auf der methodischen Ebene schwächt jedoch keinesfalls die Ergebnisse. Der Autor orientiert sich in der Auswertung zum einen an den aus der Hochschuldidaktik abgeleiteten Clustern ‚Lernerzentrierung und Dialog‘, ‚Leistungsbereitschaft‘ sowie ‚Teamarbeit in Lehre und Betreuung‘, mit denen er die Aussagen der befragten Studierenden und Lehrenden zu ihrer Haltung gegenüber dem interkulturellen Lernen beschreibt. Zum anderen



stellt er sechs praktizierte Lehrformate bzw. Betreuungskonzeptionen vor und analysiert sie im Sinne der in der Hochschuldidaktik formulierten Ziele. Als wichtiges Ergebnis seiner Studie kann festgehalten werden, dass die vorgestellten Lehrformate und Betreuungskonzepte für internationale Studierende und Lehrende einen Mehrwert für die Persönlichkeitsentwicklung, (inter-)kulturelle Sensibilität, Fähigkeit zum Perspektivwechsel, aber auch zum fachlich-disziplinären Lerngewinn erbringen. Unabdingbare Voraussetzung für diese gelingende Praxis ist die Einstellung und das Fremdverstehen der beteiligten Lehrenden und Studierenden und das Verständnis dafür, dass Internationalisierung als großes Ganzes nur dann gelingt, wenn sie im Kleinen von allen Beteiligten gewollt und getragen wird.

In seinem Resümee kommt der Autor zu dem Schluss, dass gewinnbringende hochschuldidaktische Ansätze bereits praktiziert werden, diese jedoch stärker über den eigenen fachlichen Tellerrand hinaus kommuniziert und wahrgenommen werden sollten, um eine tatsächliche Orientierung der Lehre am interkulturellen Lernen an deutschen Universitäten zu verwirklichen.

Das Buch kann als Beispiel einer gelungenen hochschuldidaktischen Begleitforschung neuer Lehrkonzepte und Betreuungsformate eingeschätzt werden; es ist eine anregende Lektüre für alle diejenigen, die interkulturelles Lernen an der Hochschule ermöglichen wollen. Notwendig erscheinen ihm eine vermehrte wissenschaftliche Begleitung innovativer hochschuldidaktischer Ansätze und eine stärkere Diskussion darüber innerhalb der scientific community, die ihrerseits hoffentlich die von Uwe Küchler am universitären Alltag empirisch überprüften Ergebnisse wahrnimmt. Die hier vorgestellte Auslandsamerikanistik kann als praktische Impulsgeberin für Konzepte der Internationalisierung und die daran anschließende wechselseitige Reflexion kultureller Phänomene in der eigenen Kultur gesehen werden.

■ **Judith Ricken & Bianca Roters**, Stipendiatinnen im Promotionskolleg „Wissensmanagement und Selbstorganisation im Kontext hochschulischer Lehr- und Lernprozesse“, Universität Dortmund

Exzellenz auch in der Lehre? - Erste Bilanztagung zur Exzellenzinitiative

Das Ende der zweiten und vorerst letzten Runde der Exzellenzinitiative des Bundes und der Länder nahm das Institut für Hochschulforschung Wittenberg zum Anlass, am 23. und 24. November eine Zwischenbilanz der nunmehr dreijährigen Geschichte der Exzellenzinitiative zu ziehen. Auf der Tagung forderten Hochschulforscher und Wissenschaftspolitiker u.a., nun Exzellenz in der Lehre mehr Geltung zu verschaffen. Sachsens Wissenschaftsministerin Eva-Maria Stange sagte: „Was wir brauchen, sind Kriterien für die Frage: Was ist gute Lehre?“

An den zwei Tagen wurden insgesamt drei Themenblöcke erörtert: Grundlagen, Praxis und Konsequenzen der Exzellenzinitiative.

Grundlagen der Exzellenzinitiative

Zum Auftakt der Tagung schilderte Edelgard Bulmahn, die ehemalige Bundesbildungsministerin, wie die Idee der Elite-Unis entstand: „Die Exzellenzinitiative: Genese einer bildungspolitischen Idee“. Was daraus wurde, so schätzt sie ein, „hat insgesamt eineinhalb Jahre Nervenkrieg gelohnt“. Auf die Freistellung der besten Forscher von der Lehre angesprochen, sagte sie jedoch: „Für eine gewisse Zeit soll dies möglich sein. Ich meine jedoch, die besten Forscher müssen auch in die Erstsemesterveranstaltungen. Dies ist wichtig für die Motivation der Studenten.“ Und sie fügte hinzu: „Eigentlich sollte von Anfang an auch die Lehre in die Exzellenzinitiative einbezogen werden.“ Dies sei aber im Zuge der Föderalismusreform ins Hintertreffen geraten. „Die Exzellenzinitiative als politisches Programm“ stellte Peer Pasternack, Forschungsdirektor des Instituts für Hochschulforschung Wittenberg, in seinem Beitrag vor. Die Frage lautete, ob diese einen Paradigmenwechsel bedeute. Seine These: Die Exzellenzinitiative sei lediglich Teil eines (auch vom Finanzvolumen) insgesamt noch viel größeren Prozesses, in dem die befristete Projektarbeit die mehr und mehr dominierende Form von Wissenschaft werde. Er unterschied hierbei drei Projektförderarten: die katalytische Förderung, die Kompaktförderung und der permanente Wettbewerb. Die katalytische Förderung bedeute eine Anschubfinanzierung bis zu einem selbsttragenden Prozess, die Kompaktförderung eine regelmäßige Anschlussfinanzierung. Bulmahn hatte zuvor die Exzellenzinitiative eher in der Art einer katalytischen Förderung vorgestellt; die Leuchttürme der deutschen Forschung sollen nach einer gewissen Zeit aus eigener Kraft weiterstrahlen.

In der anschließenden Diskussion wurde eine zwiespältige Bilanz gezogen. Während die anwesenden Universitätsrektoren und -kanzler vorrangig positive Effekte sahen, wurde andererseits auch stärker Kritik laut: Gefördert werden vor allem solche Universitäten, die bereits besser ausgestattet sind; Fachhochschulen waren nicht antragsberechtigt; naturwissenschaftliche Themen dominieren geisteswissenschaftliche; prämiert werden Leistungsversprechen statt Leistungen. In der Podiumsdiskussion zu diesem Themenblock musste auch nachdenklich stimmen, dass den meisten Universitäten kaum möglich war, nicht an der Exzellenzinitiative teilzunehmen.

„Es geht nicht um Geld, sondern um Prestige“, resümierte Peter Scharff, Rektor der TU Ilmenau.

Praxis der Exzellenzinitiative

Einen scharfen Kontrast zwischen einer Universität nach Derrida und einer aktuellen Exzellenzuniversität stellte Wolfgang Fach, Prorektor der Universität Leipzig, in seinem Beitrag „Die unbedingte Universität“ vor. Nach Derrida sei eine solche Universität u.a. durch Askese, Distanz und Machtlosigkeit nach außen gekennzeichnet. Anhand von Stellenausschreibungen der TU München in der Folge der Exzellenzinitiative verdeutlichte er den anstehenden Paradigmenwechsel: Im Zentrum der Universität der Exzellenz stünden nicht mehr Ideen, sondern Interessen, die aggressiv gegenüber den Konkurrenten durchgesetzt werden. Ziel sei immer das Extrem, nämlich die beste, frauenfreundlichste, internationalste etc. Universität zu sein.

Über „Legitimität, Transparenz, Partizipation – Praxis des Begutachtens“ bei der Auswahl der Graduiertenkollegs, Exzellenzcluster und Zukunftskonzepte referierte Michael Zürn (Hertie School of Governance). Die Wissenschaftlichkeit bei der Begutachtung und deren Transparenz sah er insgesamt als gegeben. Aber je nachdem, inwieweit der jeweilige Berichtersteller in der Lage war, die Begeisterung der Gutachter an fachfremde Entscheider zu vermitteln, hätten prinzipiell gleichrangige Anträge unterschiedliche Chancen gehabt. Hier habe er die größten Zweifel. Denn es gäbe ein Problem, die Besten aufzuzeigen, wenn z.T. die Reputation bzw. die bisher erbrachten Forschungsleistungen und z.T. die Qualität der Anträge bzw. der zu erwartenden Leistungen ausschlaggebend war. Insbesondere in den Geisteswissenschaften wurden die Anträge generell zurückhaltender begutachtet, was sich für diese nachteilig auswirkte. Um den unterschiedlichen Fächerkulturen Rechnung zu tragen, formulierte er den Vorschlag einer „weichen Quotierung“ nach den vier Wissenschaftsgebieten, wobei die Anträge erst zum Schluss zusammengeführt würden. Außerdem sollte bei einer Verstetigung der Exzellenzinitiative weitergefördert werden, wer die versprochenen Leistungen erbracht habe. Ein Teil sollte ggf. „absteigen“ – ähnlich wie bei der Fußball-Bundesliga.

In der Podiumsdiskussion ging es dann mit der sächsischen Wissenschaftsministerin um „Exzellenz – auch in der Lehre“. In ihrem Impulsreferat forderte Eva-Maria Stange, Exzellenz in der Lehre die nötige Geltung zu verschaffen. Hier war sie sich mit Vertreterinnen der Bundesvertretung Akademischer Mittelbau (Petra Jordan) und des freien Zusammenschlusses der Studierendenenschaften (Imke Buß) einig. Ihre Ankündigung, in Sachsen ein hochschuldidaktisches Zentrum zur Verbesserung der Lehre einzurichten, erhielt ebenfalls Zuspruch. Der Vorschlag von Stange, durch leistungsorientierte Mittelverteilung zur Verbesserung der Lehre beizutragen, erntete jedoch auch Widerspruch – insbesondere wegen der oft unzureichenden Grundfinanzierung.



Bulmahn fordert: „Die besten Forscher müssen auch in die Erstsemesterveranstaltungen“

Foto: König/ HoF Wittenberg

Konsequenzen der Exzellenzinitiative

Zur Frage „Was bedeutet die Exzellenzinitiative für die Nachwuchsförderung?“ referierte Johannes Moes von der GEW-Projektgruppe Doktorand/innen. Unter anderem stellte er dar, dass die bei der Exzellenzinitiative Siegreichen auch die „doctoral mills“ Deutschlands sind, indem sie 51% der abgeschlossenen Promotionen auf sich vereinigen. Dies bedeute auch eine große Chance für die Nachwuchsförderung, wenn die Förderung dort mehr als nur einigen wenigen Promovierenden zugute kommt. Wiederum in einer Podiumsdiskussion stellten dann Vertreter verschiedener Graduiertenschulen Vorzüge und Probleme der Konzepte ihrer Institution vor. Als Vorzüge wurden die z.T. in ihrer Transparenz vorbildlichen Auswahlverfahren gesehen. Als problematisch eingeschätzt wurde, dass vereinzelt bis zur Hälfte der Promotionszeit für den Besuch von Lehrveranstaltungen und Kolloquien verpflichtend festgelegt ist.

„Gibt es objektive Kriterien für Exzellenz?“ fragte Stefan Hornbostel, Leiter des Institutes für Forschungsinformation und Qualitätssicherung (IFQ), das mit der Evaluation der Exzellenzinitiative beauftragt wurde. Seine Antwort könnte man mit „Jein“ umschreiben. Für die in der Vergangenheit erbrachten Veröffentlichungen, eingeworbenen Forschungsgelder und die Qualität der Nachwuchsförderung ließen sich noch relativ gut nachvollziehbare Kriterien und Indikatoren finden, von denen er einige vorstellte. Wesentlich schwieriger sehe es mit solchen Aspekten wie Originalität aus. Anhand von ausgewählten Ergebnissen der Evaluation des Emmy-Noether-Programmes für exzellente Nachwuchswissenschaftler/innen zeigte er, dass diese nicht unbedingt häufiger publizieren oder zitiert und berufen werden. Zusammenfassend räumte er ein: „Leistungen und die Art ihrer Messung werden letztlich in Kommunikationsprozessen definiert und evaluiert. Sie sind daher Bestandteil von Aushandlungsprozessen.“ „Die Folgen der Exzellenzinitiative für die deutsche Hochschullandschaft“ waren das Thema der abschließenden Podiumsdiskussion. Anke Burkhardt, Ge-

schäftsführerin und stellvertretende Direktorin des Institutes für Hochschulforschung Wittenberg, sieht in der Exzellenzinitiative die Chance, dem Zusammenhang von Gleichstellungsfortschritt und Qualitätsgewinn in der Wissenschaft mehr Geltung zu verschaffen. Immerhin seien drei der Gewinner-Universitäten auch in der Spitzengruppe des CEWS-Gleichstellungsrankings vertreten. Als wünschenswert stellte sie eine komplette Dokumentation der Anträge und Gutachten sowie die verbindliche Einbeziehung von Gleichstellungsaspekten in die geplante Evaluation heraus. Andreas Keller, Leiter des Vorstandsbereichs Hochschule und Forschung in der GEW, forderte schließlich vor dem Hintergrund der vielen auf der Tagung angesprochenen Aspekte die kritische Beobachtung der weiteren Entwicklung und eine offene Diskussion der Folgen der Exzellenzinitiative. Als Ergebnis der Diskussionen kann insgesamt festgehalten werden: Große Einigkeit herrschte bei den Teilnehmern darin, dass die Lehre in künftigen Exzellenz-Programmen mehr Gewicht erhalten und die Vergabeverfahren weiterentwickelt werden sollten. Die meisten Teilnehmer hielten zur Sicherung der Nachhaltigkeit der eingesetzten Mittel auch eine Verstetigung der Exzellenzinitiative für sinnvoll. Allerdings dürfte dies keinesfalls nach dem Matthäus-Prinzip geschehen (Wer hat, dem wird gegeben). Vielmehr erhielten Vorschläge der Evaluation der Gewinner bei Folgeantragstellung und Prinzipien, nach denen ein bestimmter Anteil an Hochschulen nicht wiederholt Gewinner sein könne. Dies würde auch der Intention der katalytischen Förderung entsprechen. Ein Tagungsteilnehmer formulierte hierzu treffend: „Die heutigen Gewinner darf man nicht in Stein meißeln.“ Denn dies verhindere die Motivation der heutigen z.T. sehr knappen Verlierer – und hemmt auf längere Sicht auch die Weiterentwicklung der heutigen Gewinner.

Fazit

Der Gewinn der Tagung lag insbesondere in den „Insider“-Berichten, der fundierten kritischen Reflektion und der Perspektivenvielfalt an Weiterentwicklungsvorschlägen durch die Referenten und Diskussionsteilnehmer. Bedauerlich waren einige kurzfristige krankheitsbedingte Absagen u.a. vom Sprecher der Fachhochschulen in der HRK und dass viele Handzeichen von Diskussionsteilnehmern nicht in Wortbeiträge mündeten. Es war nachvollziehbar, dass aufgrund des Zeitregimes oft schon nach zwei Sammelrunden die Rednerliste geschlossen werden musste. Eventuell würden ein Weniger an Podiumsdiskussionen und längere offene Diskussionen der enormen Debattierfreude vieler Tagungsteilnehmer bei diesem Thema noch stärker entgegenkommen. Die Tagung kann sicherlich nicht als das Ende, sondern eher als Auftakt zur intensiveren wissenschaftlichen Auseinandersetzung mit der Exzellenzinitiative gelten. Hierfür lieferte sie in jedem Fall sehr viele Impulse. Die Ergebnisse der Tagung werden 2008 veröffentlicht. Die Tagung fand in Kooperation mit der Gewerkschaft Erziehung und Wissenschaft (GEW) und mit Unterstützung des Stifterverbandes für die Deutsche Wissenschaft statt.

■ Dr. René Krempkow, Institut für Hochschulforschung (HoF), Wittenberg

Hauptbeiträge der aktuellen Hefte IVI, P-OE, QiW, HM und ZBS

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

IVI

Internationalisierung, Vielfalt und Inklusion in Hochschulen

Forum für Interkulturalität, Diversity-Management, Anti-Diskriminierung und Inklusion

Heft 3/2007

Anti-Diskriminierung und Diversity

Politik und Entwicklung im Themenbereich

Ingo Püttner

Die Anwendung des Allgemeinen Gleichbehandlungsgesetzes (AGG) in der Praxis des öffentlichen Dienstes

Günther Vedder

Chancengleichheit und Diversität an deutschen Hochschulen

Renate Klees-Möller, Hiam Tarzi & Karola Wolff-Bendik

Diversity als Chance? Mentoringprogramme für Frauen mit Migrationshintergrund an der Universität Duisburg-Essen

Rolf Meinhardt

Hochqualifizierte Einwanderer – erster BA-Studiengang für eine ignorierte Elite

Anregungen für die Praxis/ Erfahrungsberichte

Andrea Ruppert & Martina Voigt

Wie lassen sich Genderaspekte in die Lehre integrieren?

Ein Modell-Modul zur Vermittlung von Verhandlungskompetenz

P-OE

Personal- und Organisationsentwicklung in Einrichtungen der Lehre und Forschung

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

Heft 3/2007

Strategien der Personalrekrutierung

Personal- und Organisationsentwicklung/-politik

Ina Voigt

Demografischer Wandel: Herausforderungen für Personalmanagement und -entwicklung

Anregungen für die Praxis/ Erfahrungsberichte

Thomas Stelzer-Rothe

Die Berufung von Professorinnen und Professoren an Fachhochschulen als strategische Aufgabe

Tobina Brinker

Schlüsselkompetenzen als Chance zur Personal- und Organisationsentwicklung in Hochschulen

Anne Brunner

Team Games – Schlüsselkompetenzen spielend trainieren.

Spiele für Seminar und Training - Folge 3

Tagungsbericht

Absolventenstudien als Instrument der Qualitätsentwicklung für Hochschulen

(Kerstin Janson)

QiW

Qualität in der Wissenschaft

Zeitschrift für Qualitätsentwicklung in Forschung, Studium und Administration

Heft 3/2007

Systemakkreditierung

Dokumentation

Kritik an Beschlüssen des Akkreditierungsrates - Offener Brief an die Wissenschaftsminister

Interview

QiW - Qualität in der Wissenschaft Interview mit dem Präsidenten der Universität Göttingen, Prof. Dr. Kurt von Figura, zur Art der Einführung der Systemakkreditierung, insbesondere zu den Beschlüssen des Akkreditierungsrates

Qualitätsentwicklung, -politik

Wolff-Dietrich Webler

System- bzw. Prozessakkreditierung in Lehre und Studium - Ausweg aus den Dilemmata der Programmakkreditierung? Teil 2 Kontroversen und Qualitätsmaßstäbe

Katharina Mallich, Isolde Kirnbauer, Michael Schmidts &

Katharina Stowasser-Bloch

Konzept zum Qualitätsmanagement in Studium und Lehre an der Medizinischen Universität Wien mittels Abbildung von Prozesslandschaften und Definition von Leistungskriterien

Gerhard Reichmann

Überprüfung der Akzeptanz einer zentralen Evaluation von Lehrveranstaltungen am Beispiel der studentischen Lehrveranstaltungsbewertung an der Universität Graz

Wertschätzung der Zeitschrift für Beratung und Studium (ZBS)

„Wenn die Qualität der Beiträge gehalten wird, kommt keine qualitätsbewusste Beratungsstelle um die Wahrnehmung dieser Publikation herum - ein Muss für Praktikerinnen und Ausbilder.“

Othmar Kürsteiner, Berufs- und Studienberatung Zürich, in seiner Rezension der ZBS in PANORAMA, Die Fachzeitschrift für Berufsberatung, Berufsbildung, Arbeitsmarkt, H. 2/07, S. 27.

HM

Hochschulmanagement

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

Heft 4/2007

Leistungsabhängige Vergütung - Erfahrungen aus der Praxis

Entwicklung von Leitungskonzepten/Leitungspolitik

Johanne Pundt, Anja Hegen, Sylvia Kaap & Katja Kohrs

Potenziale des Bildungsmarketings am Beispiel von Promotionsstudiengängen

Benedikt Hell

Hochschulzulassung ausländischer Studierender

Die Prozentrangnormierung als Alternative zum bisherigen Umrechnungsverfahren ausländischer Sekundarabschlussnoten

Ina Voigt & Volker Klein-Moddenborg
Zielvereinbarungen: Instrument strategischer Steuerung und Basis für leistungsabhängige Vergütungssysteme

Anregungen für die Praxis/ Erfahrungsberichte

Stephan Jerusel & Katrin Bachmann
Einführung der leistungsorientierten Vergütung - Die Zeit läuft ...

Erfahrungsbericht aus Informationsveranstaltungen und Beratungen zur Umsetzung des TVöD

Henning Lindhorst
Generationswechsel bei Smartcards: Schlankere Prozesse, weniger Kosten und mehr Komfort

„Das Hochschulwesen“ ist in Norwegen eine akkreditierte und für Publikationen empfohlene Zeitschrift, in der die Autoren Punkte sammeln können.

ZBS

Zeitschrift für Beratung und Studium

Handlungsfelder, Praxisbeispiele und Lösungskonzepte

Heft 4/2007

Beratung oder Service?

Beratungsentwicklung/-politik

Franz Rudolf Menne

Das „Akademische Auskunftsamt für Studien- und Berufsfragen“ der Universität Köln (1923-1938) - Entwicklung, Ausrichtung und Arbeitsalltag einer frühen Beratungseinrichtung

Anregungen aus der Praxis/ Erfahrungsberichte

Ralf Mahler

Das ServiceCenter der Leibniz Universität Hannover – Entstehungsgeschichte und Erfahrungen nach eineinhalb Jahren Praxis. Ein Zwischenbericht

Siegfried Engl

Zwei Jahre Info-Service an der Freien Universität Berlin. Eine Bilanz. Call-Center, E-Mail-Service und Info-Counter an der Schnittstelle von Dienstleistungen für Studierende und Bewerber

Isabel von Colbe-van de Vyver & Juliane Just-Nietfeld

Zu den Auswirkungen der Einrichtung einer Studienzentrale auf die Kernaufgaben der Zentralen Studienberatung: Information und Beratung

ZBS - Zeitschrift für Beratung und Studium

Das Anliegen hinter der Frage – Beratungshandeln erfordert Professionalität - Ein Interview der ZBS mit Mitarbeiterinnen des Teams Zentrale Studienberatung der Universität Göttingen

Klaus Scholle

Kommentar: Auf dem Weg zur Kundenberatung?



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

**Falk Bretschneider/Peer Pasternack:
Handwörterbuch der Hochschulreform**



Die deutsche Hochschulreformdebatte hat sich seit einiger Zeit konzeptionell erheblich verbreitert. Vor allem die Internationalisierung incl. der europäischen Hochschulintegration sowie Referenzen betriebswirtschaftlicher Art kennzeichnen diese Entwicklungen. Mit den neuen Konzepten erweitert sich auch das Reformvokabular. Manches dabei wird gut klingende Modeerscheinung bleiben, anderes sich als unverträglich mit Auftrag und Funktion von Hochschulen erweitern. Doch viele dieser konzeptionellen und begrifflichen Transfers aus anderen Handlungsfeldern werden die Hochschulreformen dauerhaft begleiten. Das sollte indes keine Einstiegshürde für die Beteiligung an den diesbezüglichen Debatten und Entwicklungen darstellen. Um den Einstieg zu erleichtern und die Orientierung dauerhaft zu verbessern, ist das „Handwörterbuch der Hochschulreform“ der richtige Begleiter. Es erläutert in kurzen und prägnanten Begriffserklärungen sowohl das neue Reformvokabular wie auch ältere Begriffe, die im Kontext der aktuellen Hochschulreformen relevant sind.

ISBN 3-937026-38-X, Bielefeld 2005,
221 Seiten, 27.70 Euro

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

Peer Pasternack:

Politik als Besuch. Ein wissenschaftspolitischer Feldreport aus Berlin

„Niemand studiert Medizin, um anschließend krank zu werden“, gibt der Politikwissenschaftler und Hochschulforscher Peer Pasternack zu Protokoll - und geht gleichwohl für zwei Jahre in die Wissenschaftspolitik. Als Staatssekretär für Wissenschaft und Forschung im Senat von Berlin hatte er sich umständehalber insbesondere mit Haushaltspolitik zu befassen. Was das für Wissenschaftspolitik bedeutet, lässt sich hier kurzweiliger, als das Thema vermuten lässt, nachlesen. Pasternack legt einen analytischen Feldreport vor, der keine Anekdotensammlung ist, in dem sich aber, so der Autor, „manches aus Gründen, die in der jeweiligen Sache selbst liegen, nur mit ironischer Färbung“ wiedergeben lässt. Wer authentische Auskünfte über die Gestaltungsmechanismen von Hochschul- und Forschungspolitik sucht, wird mit diesem Bericht ebenso bedient, wie diejenigen, die nach Zusammenhang, Wechselwirkung und Entkopplung von analytischem Wissen und politischen Handeln fragen, wie auch diejenigen, die sich für die Untiefen der Berliner Landespolitik interessieren

ISBN 3-937026-40-1,
Bielefeld 2005, 2. Auflage, 253 Seiten, 29.70 Euro



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